SCIENCE

Science

SECTION B

A UNITED STATES DEPARTMENT OF COMMERCE PUBLICATION


# JOURNAL OF RESEARCH

of the NATIONAL BUREAU OF STANDARDS

MAY 17 1972

Volume 75B Nos. 3&4 July-Dec. 1971


# Mathematical Sciences

MAY 15 1972

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards

**Journal of Research**

**of the**

**National Bureau of Standards**

**B. Mathematical Sciences**

July–Dec. • 1971

VOLUME • 75B

NUMBERS • 3&4

Editor: Morris Newman

Associate Editor: F. W. Olver

Publication dates: Mar. 10, June 10, Sept. 10, Dec. 10, 1971

J. Res. Nat. Bur. Stand. (U.S.), B CODEN: JNBBA 75B(3&4) 97-196 (1971)

**JOURNAL OF RESEARCH**

The National Bureau of Standards Journal of Research reports research and development in the fields of activity shown at right. Also included from time to time are survey articles on topics closely related to the Bureau's scientific and technical program.

The Journal is published in three separate sections as follows:

**A. Physics and Chemistry**

Contains papers of interest primarily to scientists working in these fields. Issued six times a year. Annual subscription: domestic, $9.50; foreign, $11.75. Single copy, varies.

**B. Mathematical Sciences**

Presents studies and compilations designed mainly for the mathematician and the theoretical physicist. Issued quarterly. Annual subscription: domestic, $5.00; foreign, $6.25. Single copy $1.25.

**C. Engineering and Instrumentation**

Reports research and development results of interest chiefly to the engineer and the applied scientist. Issued quarterly. Annual subscription: domestic, $5.00; foreign, $6.25. Single copy $1.25.

The National Bureau of Standards[1] was established by an act of Congress March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau consists of the Institute for Basic Standards, the Institute for Materials Research, the Institute for Applied Technology, the Center for Computer Sciences and Technology, and the Office for Information Programs.

**THE INSTITUTE FOR BASIC STANDARDS** provides the central basis within the United States of a complete and consistent system of physical measurement; coordinates that system with measurement systems of other nations; and furnishes essential services leading to accurate and uniform physical measurements throughout the Nation's scientific community, industry, and commerce. The Institute consists of a Center for Radiation Research, an Office of Measurement Services and the following divisions:

Applied Mathematics—Electricity—Heat—Mechanics—Optical Physics—Linac Radiation[2]—Nuclear Radiation[2]—Applied Radiation[2]—Quantum Electronics[3]—Electromagnetics[3]—Time and Frequency[3]—Laboratory Astrophysics[3]—Cryogenics[3].

**THE INSTITUTE FOR MATERIALS RESEARCH** conducts materials research leading to improved methods of measurement, standards, and data on the properties of well-characterized materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; and develops, produces, and distributes standard reference materials. The Institute consists of the Office of Standard Reference Materials and the following divisions:

Analytical Chemistry—Polymers—Metallurgy—Inorganic Materials—Reactor Radiation—Physical Chemistry.

**THE INSTITUTE FOR APPLIED TECHNOLOGY** provides technical services to promote the use of available technology and to facilitate technological innovation in industry and Government; cooperates with public and private organizations leading to the development of technological standards (including mandatory safety standards), codes and methods of test; and provides technical advice and services to Government agencies upon request. The Institute also monitors NBS engineering standards activities and provides liaison between NBS and national and international engineering standards bodies. The Institute consists of the following technical divisions and offices:

Engineering Standards Services—Weights and Measures—Flammable Fabrics—Invention and Innovation—Vehicle Systems Research—Product Evaluation Technology—Building Research—Electronic Technology—Technical Analysis—Measurement Engineering.

**THE CENTER FOR COMPUTER SCIENCES AND TECHNOLOGY** conducts research and provides technical services designed to aid Government agencies in improving cost effectiveness in the conduct of their programs through the selection, acquisition, and effective utilization of automatic data processing equipment; and serves as the principal focus within the executive branch for the development of Federal standards for automatic data processing equipment, techniques, and computer languages. The Center consists of the following offices and divisions:

Information Processing Standards—Computer Information—Computer Services—Systems Development—Information Processing Technology.

**THE OFFICE FOR INFORMATION PROGRAMS** promotes optimum dissemination and accessibility of scientific information generated within NBS and other agencies of the Federal Government; promotes the development of the National Standard Reference Data System and a system of information analysis centers dealing with the broader aspects of the National Measurement System; provides appropriate services to ensure that the NBS staff has optimum accessibility to the scientific information of the world, and directs the public information activities of the Bureau. The Office consists of the following organizational units:

Office of Standard Reference Data—Office of Technical Information and Publications—Library—Office of Public Information—Office of International Relations.

[1] Headquarters and Laboratories at Gaithersburg, Maryland, unless otherwise noted; mailing address Washington, D.C. 20234.
[2] Part of the Center for Radiation Research.
[3] Located at Boulder, Colorado 80302.

U.S. DEPARTMENT OF COMMERCE
Maurice H. Stans, Secretary
NATIONAL BUREAU OF STANDARDS
Lewis M. Branscomb, Director

Order all publications from the Superintendent of Documents
U.S. Government Printing Office, Washington, D.C. 20402.

Use of funds for printing this publication approved by the Director of the Bureau of the Budget (June 22, 1966).

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# Some Elementary Formulas in "Matrix Calculus" and Their Applications

**Jeffrey T. Fong**

**Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234**

**(March 8, 1971)**

A collection of elementary formulas for calculating the gradients of scalar- and matrix-valued functions of one matrix argument is presented. Using some of the well-known properties of the operator "trace" on square matrices, alternative definitions of gradients and simple examples of calculating them using the product rule and the chain rule for differentiation are treated in an expository fashion in both component and matrix notations with emphasis on the latter. Two examples in continuum mechanics are presented to illustrate the application of the so-called "matrix calculus" of differentiable functions.

Key words: Chain rule; continuum mechanics; gradient; matrices; matrix calculus; partial differentiation; product rule; tensor function; trace.

## 1. Introduction

This is an expository article on the use of matrix notation in the elementary calculus of differentiable functions whose arguments are square matrices. For example, in continuum physics, it is often necessary to work with partial derivatives of a class of functions whose arguments are elements of a square matrix and whose values can be either scalars or square matrices of the same order. Following the notation and basic concepts of tensor functions as treated by Truesdell and Noll [1, pp. 20–35][1], we present here an elementary introduction to the proper formulation of the chain rule and the product rule for differentiation in matrix notation and we include examples, formulas and applications to illustrate the two rules.

The reader is assumed to be familiar with the notions of the trace and the determinant of a matrix $\underline{\underline{A}} = (A_{ij})$, $i, j = 1, 2, \ldots, n$, i.e., $\operatorname{tr} \underline{\underline{A}} = \sum_{i=1}^{n} A_{ii}$ and $\det \underline{\underline{A}} = \Sigma (-1)^h A_{1\sigma_1} A_{2\sigma_2} \ldots A_{n\sigma_n}$, where the last summation is made over all permutations of $\sigma_1, \sigma_2, \ldots, \sigma_n$, and $h$ is the number of interchanges required to restore the natural order.[2] In particular, the following properties of the operator "trace" are applied frequently throughout the paper:

(a) $\operatorname{tr} (\underline{\underline{A}} + \underline{\underline{B}}) = \operatorname{tr} \underline{\underline{A}} + \operatorname{tr} \underline{\underline{B}}$;

(b) $\operatorname{tr} (\underline{\underline{A}}\,\underline{\underline{B}}) = \operatorname{tr} (\underline{\underline{B}}\,\underline{\underline{A}})$;

(c) $\operatorname{tr} (\underline{\underline{A}}^T) = \operatorname{tr} \underline{\underline{A}}$, where $\underline{\underline{A}}^T$ denotes the transpose of $\underline{\underline{A}}$;

(d) $\underline{\underline{A}} = \underline{\underline{B}}$, if, and only if, $\operatorname{tr} (\underline{\underline{A}}\,\underline{\underline{C}}) = \operatorname{tr} (\underline{\underline{B}}\,\underline{\underline{C}})$ for arbitrary matrix $\underline{\underline{C}}$.

---

*AMS Subject Classification:* Primary 15, 88; Secondary 69.

[1] Figures in brackets indicate the literature references at the end of this paper.

[2] A square matrix is denoted by a symbol underlined with two bars indicating the need for two indices in component notation. In general, any quantity with, say, $k$ indices in component notation will be underlined with $k$ bars when the indices are suppressed. For ease of printing, this convention is followed in equations but ignored in text.

## 2. Gradient of a Scalar Function of a Matrix Argument

Let $\epsilon \equiv \hat{\epsilon}(A_{11}, A_{12}, \ldots, A_{1n}, A_{21}, A_{22}, \ldots, A_{2n}, \ldots, A_{n1}, A_{n2}, \ldots, A_{nn})$ define a scalar-valued function $\hat{\epsilon}$ of $n^2$ variables $A_{km}$, $k, m = 1, 2, \ldots, n$, such that the set of variables $A_{km}$ corresponds to the set of components of a square matrix $\underline{\underline{A}}$ of order $n$. In matrix notation, the definition of the scalar function $\hat{\epsilon}$ assumes the following simple form:

$$\epsilon \equiv \hat{\epsilon}(\underline{\underline{A}}). \tag{2.1}$$

If $\hat{\epsilon}$ is differentiable with respect to each variable $A_{km}$, the set of first partial derivatives of $\hat{\epsilon}$, i.e., $\{D_{km}\hat{\epsilon}, k, m = 1, 2, \ldots, n\}$, can be defined as a matrix-valued function to be denoted by $\nabla\hat{\epsilon}$ where the element $(\nabla\hat{\epsilon})_{km}$ at the $k$th row and the $m$th column of $\nabla\hat{\epsilon}$ is given precisely by $D_{km}\hat{\epsilon}$. Let $\underline{\underline{\epsilon}}_A$ denote the value of the function $\nabla\hat{\epsilon}$ for a given $\underline{\underline{A}}$, then the definition of the function $\nabla\hat{\epsilon}$, to be known as the gradient of $\hat{\epsilon}$, can be stated in both component and matrix notations as follows:

$$\underline{\underline{\epsilon}}_A \equiv [(\epsilon_A)_{km}] \equiv [D_{km}\hat{\epsilon}(A_{pq})] \equiv [\nabla\hat{\epsilon})_{km}(A_{pq})] \equiv \underline{\underline{\nabla\hat{\epsilon}}}(\underline{\underline{A}}). \tag{2.2}$$

For brevity, we omit the statement that all indices $k, m, p, q, \ldots$, etc., range from 1 to $n$.

For the purpose of applying those properties of the operator "trace" as listed in the last section, Truesdell and Noll [1] presented an alternative definition of the gradient of a scalar function of a matrix argument as follows:

$$\operatorname{tr}\{\underline{\underline{\epsilon}}_A^T \underline{\underline{C}}\} \equiv \frac{d}{ds}\hat{\epsilon}(\underline{\underline{A}} + s\underline{\underline{C}}) \mid_{s=0}, \tag{2.3a}$$

or, in component notation,

$$\sum_{k=1}^{n}\sum_{m=1}^{n}(\epsilon_A)_{km}C_{km} \equiv \frac{d}{ds}\hat{\epsilon}(A_{pq} + sC_{pq}) \mid_{s=0}, \tag{2.3b}$$

where $C$, with components $C_{pq}$, is an arbitrary matrix of the same order as the matrix $A$. To see that (2.2) and (2.3) are equivalent, we apply the chain rule for differentiation to the expression $\hat{\epsilon}(A_{pq} + sC_{pq})$:

$$\frac{d}{ds}\hat{\epsilon}(A_{pq} + sC_{pq}) = \sum_{k=1}^{n}\sum_{m=1}^{n}[D_{km}\hat{\epsilon}(A_{pq} + sC_{pq})]\left[\frac{d}{ds}(A_{km} + sC_{km})\right]$$

$$= \sum_{k=1}^{n}\sum_{m=1}^{n}\left[D_{km}\hat{\epsilon}\ (A_{pq} + sC_{pq})\right]C_{km}. \tag{2.4}$$

If we substitute zero for $s$ in (2.4) and apply (2.2), we obtain (2.3b). Conversely, (2.3b) and the chain rule imply (2.2). The reader may wish to verify that (2.3) indeed defines a unique matrix $\underline{\underline{\epsilon}}_A$ as a result of the linearity of the operator "trace" and the arbitrariness of the matrix $C$.

EXAMPLE 1: $\epsilon = \hat{\epsilon}(\underline{\underline{A}}) = \det \underline{\underline{A}}$.

To calculate the gradient of $\hat{\epsilon}$, we apply the Laplace development of a determinant, i.e.,

$$\det \underline{\underline{A}} = \sum_{k=1}^{n} A_{km}A^{km}, \ m \text{ being fixed and not summed}, \tag{2.5}$$

where $A^{km}$ denotes the cofactor of $A_{km}$ defined as $(-1)^{k+m}$ times the complementary minor of $A_{km}$. Using (2.2) and the fact that the cofactor $A^{km}$ is independent of the element $A_{km}$ of the matrix $A$, we obtain:

$$\underline{\underline{\epsilon}}_A = [(\epsilon_A)_{km}] = [D_{km}(\det \underline{\underline{A}})] = [A^{km}] = \underline{\underline{A}}^{\text{cof}} \tag{2.6}$$

where $A^{\text{cof}}$ denotes the cofactor matrix of $A$ which, by definition, equals the transpose of the adjoint matrix of $A$. Let us verify the result given in (2.6) by applying the alternative definition of $\epsilon_A$ as given in (2.3):

$$\begin{aligned} \text{tr}\{\underline{\underline{\epsilon}}_A^T \underline{\underline{C}}\} &= \frac{d}{ds} \det(\underline{\underline{A}} + s\underline{\underline{C}})\,|_{s=0} \\ &= \frac{d}{ds} \det[\underline{\underline{A}}(\underline{\underline{1}} + s\underline{\underline{A}}^{-1}\underline{\underline{C}})]\,|_{s=0} \\ &= (\det \underline{\underline{A}}) \frac{d}{ds} \det(\underline{\underline{1}} + s\underline{\underline{A}}^{-1}\underline{\underline{C}})\,|_{s=0}. \end{aligned} \tag{2.7}$$

Following Truesdell and Noll [1], we introduce another expansion of a determinant:

$$\det(1 + s\underline{\underline{B}}) = 1 + I_1(\underline{\underline{B}})s + I_2(\underline{\underline{B}})s^2 + \ldots + I_n(\underline{\underline{B}})s^n, \tag{2.8}$$

where $B$ is any square matrix of order $n$ and $I_1(B), I_2(B), \ldots, I_n(B)$ are the so-called principal invariants of $B$.[3] In our case, we are only interested in the first principal invariant $I_1(B)$ which equals tr $(B)$. Combining (2.7) and (2.8), we obtain:

$$\begin{aligned} \text{tr}\{\underline{\underline{\epsilon}}_A^T \underline{\underline{C}}\} &= (\det \underline{\underline{A}})\ \text{tr}(\underline{\underline{A}}^{-1}\underline{\underline{C}}) \\ &= \text{tr}\{(\det \underline{\underline{A}})\underline{\underline{A}}^{-1}\underline{\underline{C}}\}, \end{aligned}$$

i.e.,

$$\underline{\underline{\epsilon}}_A = (\det \underline{\underline{A}})(\underline{\underline{A}}^{-1})^T. \tag{2.9}$$

Since $\underline{\underline{A}}^{-1} = (\det \underline{\underline{A}})^{-1}\ (\underline{\underline{A}}^{\text{cof}})^T$, we see immediately that (2.9) is equivalent to (2.6), and that both definitions given in (2.2) and (2.3) yield the same result.

EXAMPLE 2: $\epsilon = \hat{\epsilon}(\underline{\underline{A}}) = \text{tr}(\underline{\underline{A}}^m)$, $m$ being any positive integer.

Since we have yet to introduce the notion of the gradient of a matrix-valued function, we must rule out the possibility of calculating the gradient of $\hat{\epsilon}$ using the chain rule. To apply the definition of the gradient of $\hat{\epsilon}$ as given in (2.2), it is necessary to develop an expansion of the function $\hat{\epsilon}$ in terms of the components of $A$. We observe that for arbitrary positive integer $m$, the expansion of the matrix $A^m$ is cumbersome, and it is not practical to find the gradient of $\hat{\epsilon}$ using (2.2).

However, the definition given in (2.3) does lead us to an answer:

$$\begin{aligned} \text{tr}\{\underline{\underline{\epsilon}}_A^T \underline{\underline{C}}\} &= \frac{d}{ds}\ \text{tr}\{(\underline{\underline{A}} + s\underline{\underline{C}})^m\}_{s=0} \\ &= \frac{d}{ds}\ \{(\underline{\underline{A}}^m + (\underline{\underline{A}}^{m-1}\underline{\underline{C}} + \underline{\underline{A}}^{m-2}\underline{\underline{C}}\underline{\underline{A}} + \ldots + \underline{\underline{C}}\underline{\underline{A}}^{m-1})s + \ldots)\}_{s=0} \\ &= \text{tr}\{\underline{\underline{A}}^{m-1}\underline{\underline{C}} + \underline{\underline{A}}^{m-2}\underline{\underline{C}}\underline{\underline{A}} + \ldots + \underline{\underline{C}}\underline{\underline{A}}^{m-1}\} \\ &= \text{tr}\{m\underline{\underline{A}}^{m-1}\underline{\underline{C}}\}, \end{aligned} \tag{2.10}$$

[3] For a rigorous exposition of the notion of a principal invariant of a matrix or a second order tensor, see Ericksen [2, p. 832].

since the trace operator is linear and $\mathrm{tr}(\underline{\underline{A}}\,\underline{\underline{B}}) = \mathrm{tr}(\underline{\underline{B}}\,\underline{\underline{A}})$. Since $\underline{\underline{C}}$ *is arbitrary, we conclude that* $\underline{\underline{\epsilon_A}} = m(\underline{\underline{A}}^{m-1})^T$.

EXAMPLE 3: $\epsilon = \hat{\epsilon}(\underline{\underline{A}}) = \det(\underline{\underline{A}}^2 + \underline{\underline{B}})$, $\underline{\underline{B}}$ being a constant matrix.

In this case, both definitions given in (2.2) and (2.3) are not practical for us to evaluate the gradient of the scalar function $\epsilon$. The only reasonable alternative is to use the chain rule in conjunction with a practical way of evaluating the gradient of a matrix-valued function as to be presented in the next section.

## 3. Gradient of a Matrix-Valued Function of a Matrix Argument

Let $f \equiv \hat{f}(A)$ define a matrix-valued function $\hat{f}$ of a matrix argument $A$ where both $f$ and $A$ are square matrices of order $n$ with components $f_{km}$ and $A_{rs}$ respectively, and the $n^2$ component functions $\hat{f}_{km}$ of $\hat{f}$ are defined as follows:

$$f_{km} \equiv \hat{f}_{km}(A_{rs}) \equiv \hat{f}_{km}(\underline{\underline{A}}). \tag{3.1}$$

If each component function $\hat{f}_{km}$ is differentiable, the set of the first partial derivatives of $\hat{f}_{km}$, i.e., $\{D_{pq}\hat{f}_{km}\}$, can be defined as the gradient of the function $\hat{f}$ to be denoted by $\nabla\hat{f}$. To emphasize the need for four indices to specify $f_A$ which stands for the value of $\nabla\hat{f}$ for a given $A$, we introduce the unusual four-bar notation as it appears in the following definition:

$$\underline{\underline{f_A}} \equiv [(f_A)_{kmpq}] \equiv [D_{pq}\hat{f}_{km}(A_{rs})] = [(\nabla\hat{f})_{kmpq}(A_{rs})] \equiv \underline{\underline{\nabla\hat{f}}}\ (\underline{\underline{A}}). \tag{3.2}$$

Clearly $f_A$ is not a square matrix in the usual sense, and, therefore, is not suitable for calculations in matrix notation. Following [1], we introduce the so-called contraction operation on $f_A$ with respect to an arbitrary square matrix $C$ whose order is the same as that of $A$:

$$\underline{\underline{f_A[C]}} \equiv \left[\sum_{p=1}^{n}\sum_{q=1}^{n}(f_A)_{kmpq}C_{pq}\right]. \tag{3.3}$$

The new quantity, $f_A[C]$, to be known as the gradient of $\hat{f}$ with respect to $A$ and contracted with $C$, requires only two indices for component representation. Hence the symbol $f_A[C]$ will replace $f_A$ wherever matrix operations are used.

The definition of the gradient of $\hat{f}$ as given in (3.2) is equivalent to the following alternative definition based on the chain rule:

$$\underline{\underline{f_A[C]}} = \frac{d}{ds}\underline{\underline{\hat{f}}}(\underline{\underline{A}} + s\underline{\underline{C}}), \tag{3.4a}$$

or, in component notation,

$$\sum_{p=1}^{n}\sum_{q=1}^{n}(f_A)_{kmpq}C_{pq} = \frac{d}{ds}\hat{f}_{km}(A_{rt} + sC_{rt}). \tag{3.4b}$$

As a rule, both definitions given in (3.2) and (3.4) are useful for simple matrix-valued functions such as those listed below:

$$\underline{\underline{\hat{f}}}(\underline{\underline{A}}) = \underline{\underline{A}}; \qquad \underline{\underline{f_A[C]}} = \underline{\underline{C}}. \tag{3.5}$$

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{A}}^T; \qquad \underline{\underline{f_A[C]}} = \underline{\underline{C}}^T. \tag{3.6}$$

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{B}}\,\underline{\underline{A}}, \; B \text{ being a constant matrix}; \quad \underline{\underline{f_A[C]}} = \underline{\underline{B}}\,\underline{\underline{C}}. \tag{3.7}$$

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{A}}\,\underline{\underline{B}}, \; B \text{ being a constant matrix}, \quad \underline{\underline{f_A[C]}} = \underline{\underline{C}}\,\underline{\underline{B}}. \tag{3.8}$$

For moderately complicated matrix-valued functions such as $\hat{f}(A) = A^m$, $m$ being a positive integer greater than 1, the matrix definition given by (3.4a) is far superior and sometimes becomes the sole means of evaluating the gradient of a matrix-valued function. The reader can easily verify, using (3.4a), the following useful result: (Note: $\underline{\underline{A}}^0 \equiv \underline{\underline{1}}$).

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{A}}^m, \; m = 2,3,4, \ldots; \qquad \underline{\underline{f_A[C]}} = \sum_{i=0}^{m-1} \underline{\underline{A}}^i \underline{\underline{C}} \underline{\underline{A}}^{m-i-1}. \tag{3.9}$$

## 4. Product Rule for Differentiation of Matrix-valued Functions

Let $\hat{f}$ be the product of two matrix-valued functions $\hat{g}$ and $\hat{h}$ with $f \equiv \hat{f}(A) \equiv \hat{g}(A)\hat{h}(A) \equiv gh$, where the order of the matrix multiplication is important. The product rule for partial differentiation yields the gradient of $\hat{f}$ in the following matrix notation:

$$\underline{\underline{f_A[C]}} = \underline{\underline{g_A[C]}}\,\underline{\underline{h}} + \underline{\underline{g}}\,\underline{\underline{h_A[C]}} \qquad \text{for} \qquad \underline{\underline{f}} = \underline{\underline{g}}\,\underline{\underline{h}} \tag{4.1}$$

Using the elementary formulas given in (3.5) and (3.6), we obtain immediately the following formula based on (4.1):

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{A}}^T\underline{\underline{A}}; \qquad \underline{\underline{f_A[C]}} = \underline{\underline{C}}^T\underline{\underline{A}} + \underline{\underline{A}}^T\underline{\underline{C}}. \tag{4.2}$$

To derive the formula for the gradient of the matrix inversion operator, we apply the product rule to the identity $\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{1}}$:

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{A}}^{-1}; \qquad \underline{\underline{f_A[C]}}\,\underline{\underline{A}} + \underline{\underline{A}}^{-1}\underline{\underline{C}} = \underline{\underline{O}}, \text{ i.e.,} \qquad \underline{\underline{f_A[C]}} = -\underline{\underline{A}}^{-1}\underline{\underline{C}}\,\underline{\underline{A}}^{-1}. \tag{4.3}$$

Using the product rule and (4.3), the reader can easily verify by induction:

$$\hat{f}(\underline{\underline{A}}) = \underline{\underline{A}}^{-m}, \; m = 2,3, \ldots; \qquad \underline{\underline{f_A[C]}} = -\sum_{i=0}^{m-1} \underline{\underline{A}}^{-m+i}\,\underline{\underline{C}}\,\underline{\underline{A}}^{-1+i}. \tag{4.4}$$

Whenever the inverse of a matrix is mentioned, the restriction to the class of square matrices with nonzero determinants will be understood.

## 5. Chain Rule for Differentiation of Scalar- and Matrix-Valued Functions

Consider a scalar-valued function $\hat{\epsilon}$ of a matrix argument $\underline{\underline{A}}$ whose components $A_{km}$ are functions of a single scalar parameter $t$. The chain rule for differentiation with respect to $t$ assumes the following form in component notation:

$$\eta(t) \equiv \hat{\epsilon}(A_{km}(t)); \qquad \dot{\eta}(t) = \sum_{k=1}^{n} \sum_{m=1}^{n} (\epsilon_A)_{km} \dot{A}_{km}(t), \tag{5.1a}$$

where the dot symbol denotes the operator $\frac{d}{dt}$. In matrix notation, (5.1a) becomes

$$\eta(t) \equiv \hat{\epsilon}(\underline{\underline{A}}(t)); \qquad \dot{\eta}(t) = \mathrm{tr}(\underline{\underline{\epsilon}}_A^T \dot{\underline{\underline{A}}}\ (t)). \tag{5.1b}$$

In applications, it is common to work with a scalar-valued function $\hat{\phi}$ of a matrix argument $f$ which depends on another matrix argument $A$. The chain rule for differentiation with respect to $A$ can be written in the following component notation:

$$\hat{\epsilon}(A_{rt}) \equiv \hat{\phi}(\hat{f}_{uv}(A_{ij})); \qquad (\epsilon_A)_{pq} = \sum_{k=1}^{n} \sum_{m=1}^{n} (\phi_f)_{km} (f_A)_{kmpq}. \tag{5.2a}$$

To write (5.2a) in matrix notation, let us contract both sides of (5.2a) with an arbitrary matrix $C$;

$$\hat{\epsilon}(\underline{\underline{A}}) \equiv \hat{\underline{\underline{\phi}}}\ (\hat{\underline{\underline{f}}}\ (A)); \qquad \mathrm{tr}(\underline{\underline{\epsilon}}_A^T\ \underline{\underline{C}}) = \mathrm{tr}(\underline{\underline{\underline{\underline{\phi}}}}_f^T\ \underline{\underline{\underline{\underline{f}}}}_A[C]). \tag{5.2b}$$

Returning to Example 3 given in section 2, we are now equipped to evaluate the gradient of the function $\hat{\epsilon}$ defined by $\epsilon = \hat{\epsilon}(\underline{\underline{A}}) = \det(\underline{\underline{A}}^2 + \underline{\underline{B}})$, $B$ being a constant matrix. Using (2.9), (3.9) and the chain rule given by (5.2b), we have:

$$\mathrm{tr}\{\underline{\underline{\epsilon}}_A^T \cdot \underline{\underline{C}}\} = \mathrm{tr}\{\det(\underline{\underline{A}}^2 + \underline{\underline{B}})(\underline{\underline{A}}^2 + \underline{\underline{B}})^{-1}(\underline{\underline{A}}\,\underline{\underline{C}} + \underline{\underline{C}}\,\underline{\underline{A}})\}.$$

Since $\underline{\underline{C}}$ is arbitrary, we have $\underline{\underline{\epsilon}}_A^T = \det(\underline{\underline{A}}^2 + \underline{\underline{B}})[(\underline{\underline{A}}^2 + \underline{\underline{B}})^{-1}\underline{\underline{A}} + \underline{\underline{A}}(\underline{\underline{A}}^2 + \underline{\underline{B}})^{-1}]$.

## 6. A Collection of Some Elementary Formulas in "Matrix Calculus"

Based on the product rule and the chain rule for differentiation in matrix notation as presented in the last two sections, a calculus of differentiable functions of square matrices, to be referred

| Formula number | Function | Gradient | Remark |
|---|---|---|---|
| 1 | $f(A) = A$ | $f_A[C] = C$ | |
| 2 | $\hat{f}(A) = A^T$ | $C^T$ | |
| 3 | $\hat{f}(A) = BA$ | $BC$ | $B$ being a constant matrix. |
| 4 | $\hat{f}(A) = AB$ | $CB$ | $B$ being a constant matrix. |
| 5 | $\hat{f}(A) = A^M$ | $\sum_{i=0}^{m-1} A^i C\, A^{m-i-1}$ | $m = 2, 3, 4, \ldots$ |
| 6 | $\hat{f}(A) = A^T A$ | $C^T A + A^T C$ | |
| 7 | $\hat{f}(A) = A^{-1}$ | $-A^{-1} C A^{-1}$ | |
| 8 | $\hat{f}(A) = A^{-m}$ | $-\sum_{i=0}^{m=1} A^{-m+i} C\, A^{-1+i}$ | $m = 2, 3, 4, \ldots$ |
| 9 | $\hat{\epsilon}(A) = \mathrm{tr}\, A$ | $\epsilon_A = 1$ | First invariant. |
| 10 | $\hat{\epsilon}(A) = \det A$ | $\det A\ (A^{-1})^T$ | $n$th invariant. |
| 11 | $\hat{\epsilon}(A) = \phi(A^T)$ | $(\phi_{A^T})^T$ | |
| 12 | $\hat{\epsilon}(A) = \hat{\phi}(BAD)$ | $B^T (\phi_{BAD})\, D^T$ | $B$, $D$ being constant matrices. |
| 13 | $\hat{\epsilon}(A) = \hat{\phi}(A^{-1})$ | $-(A^{-1})^T (\phi_{A^{-1}})(A^{-1})^T$ | |
| 14 | $\hat{\epsilon}(A) = \hat{\phi}(AA^T)$ | $2(\phi_{AA^T})A$ | Note $(\phi_{AA^T})^T = (\phi_{AA^T})$. |
| 15 | $\hat{\epsilon}(A) = \hat{\phi}(A^T A)$ | $2A(\phi_{A^T A})$ | Note $(\phi_{A^T_{\ T_A}})^T = (\phi_{A^T_{\ T_A}})$. |
| 16 | $\hat{\epsilon}(A) = \mathrm{tr}\, A^{-1} \det A$ | $\det A\ (A^{-1})^T (1\ \mathrm{tr}\, A^{-1} - (A^{-1})^T)$ | Second invariant for $n = 3$. |

to as the "matrix calculus," can be defined in an analogous way as the elementary theory of calculus based on the field of real or complex numbers. Obviously, for matrix calculus, the underlying mathematical object is not a field, but a noncommutative ring, i.e., the ring of square matrices of order $n$ over the familiar ring of differentiable functions. An excellent account of the theory of matrices over rings was given recently by Newman [3], but here we merely present a collection of some elementary formulas in "matrix calculus" without studying its mathematical structure. For the convenience of the reader, the table on page 102 lists some of the most commonly used formulas in matrix calculus.

Using the properties of the operator "trace" as listed in section 1, we observe that the derivation for formulas Nos. 11–15 presents no difficulty. For example, formula No. 13 can be derived as follows:

From (5.2b) and formula 7, we have

$$\operatorname{tr}\,(\underline{\underline{\epsilon}}_A^T\,\underline{\underline{C}}) = \operatorname{tr}\,(-\underline{\underline{\underline{\underline{\phi}}}}_{A^{-1}})^T\underline{\underline{A}}^{-1}\underline{\underline{C}}\underline{\underline{A}}^{-1}) = \operatorname{tr}\,(-\underline{\underline{A}}^{-1}\underline{\underline{\underline{\underline{\phi}}}}_{A^{-1}})^T\underline{\underline{A}}^{-1}\underline{\underline{C}}). \tag{6.1}$$

Since $C$ is arbitrary, we obtain immediately the desired result.

## 7. Applications

To illustrate the ease with which certain problems in continuum physics can be treated by using some of the formulas listed in the last section, we shall present two examples in continuum mechanics:

EXAMPLE 1: (All indices $i, j, k, m, p$, etc. range from 1 to 3.)

Let the material coordinates of a particle in a continuous body be denoted by $X^i$. Let the position coordinates of the same particle at time $t$ be given by $x^k \equiv \hat{x}^k(X^i, t)$. Two basic quantities can be defined:

$$\underline{\underline{F}} \equiv \hat{F}^k_I(X^j, t) \equiv \frac{\partial \hat{x}^k}{\partial X^i}(X^j, t); \qquad \text{(deformation gradient);} \tag{7.1}$$

$$\underline{v} \equiv \hat{v}^k(X^j, t) \equiv \frac{\partial \hat{x}^k}{\partial t}(X^j, t); \qquad \text{(velocity vector).} \tag{7.2}$$

It is useful to express $X^i$ as functions of $x^k$ and $t$ so that the velocity components have the alternative representation $v^k = \tilde{v}^k(x^m, t)$. This allows us to define another useful quantity:

$$\underline{\underline{L}} \equiv \hat{L}^k_m(x^p, t) \equiv \frac{\partial \tilde{v}^k}{\partial x^m}(x^p, t), \qquad \text{(velocity gradient).} \tag{7.3}$$

An important relation to be needed later follows immediately from the above definitions and the interchange of the order of partial differentiations:

$$\dot{\underline{\underline{F}}} = \underline{\underline{L}}^T\underline{\underline{F}}, \tag{7.4}$$

where the dot symbol denotes the partial derivative with respect to $t$ holding the material coordinates $X^i$ constant.

The notion of "mass" of a continuous body leads to two notions of "mass density", namely,

the mass density $\rho_R$ with respect to a unit volume in the reference configuration where each particle is labeled with the material coordinates $X^i$, and the mass density $\rho_t$ with respect to a unit volume in the spatial configuration at time $t$ where each particle is observed to occupy the position at coordinates $x^k$. The two mass densities are of course related:

$$\rho_R = \rho_t \det \underline{\underline{F}}. \tag{7.5}$$

The law of the conservation of mass states that $\dot{\rho}_R = 0$. Using (5.1b), formula No. 10, and the relation (7.4), we obtain the well known "equation of continuity" in classical mechanics: (Note: $\operatorname{div} \underline{v} \equiv \sum_{k=1}^{3} \frac{\partial v^k}{\partial x^k}$.)

$$\begin{aligned}
\dot{\rho}_R &= \dot{\rho}_t \det \underline{\underline{F}} + \rho_t \dot{(\det \underline{\underline{F}})} \\
&= \dot{\rho}_t \det \underline{\underline{F}} + \rho_t \operatorname{tr} (\det \underline{\underline{F}} (\underline{\underline{F}}^{-1}) \dot{\underline{\underline{F}}}) \\
&= \dot{\rho}_t \det \underline{\underline{F}} + \rho_t \det \underline{\underline{F}} \operatorname{tr} (\underline{\underline{F}}^{-1} \underline{\underline{L}}^T \underline{\underline{F}}) \\
&= \dot{\rho}_t \det \underline{\underline{F}} + \rho_t \det \underline{\underline{F}} \operatorname{tr} (\underline{\underline{L}}^T) \\
&= \det \underline{\underline{F}} (\dot{\rho}_t + \rho_t \operatorname{div} \underline{v})
\end{aligned}$$

Since $\det \underline{\underline{F}} \neq 0$, $\dot{\rho}_R = 0$ implies $\dot{\rho}_t + \rho_t \operatorname{div} \underline{v} = 0$.

EXAMPLE 2:

One of the principles generally associated with the correct formulation of the constitutive equation of a material is known as the "Principle of Material Indifference" which means physically that the response of a material is independent of the observer. Let us confine our attention to "simple fluids" in the sense of Truesdell and Noll [1], where the most general constitutive equation may be written in the following form:

$$\underline{\underline{T}}(t) = \underline{\underline{\mathscr{G}}}[\underline{\underline{C}}_t(\tau);\ \rho_t]. \tag{7.6}$$

Here $T(t)$ is the Cauchy stress at time $t$, $C_t(\tau)$ is the relative right Cauchy-Green tensor defined by $(C_t(\tau))_{km} \equiv \sum_{p=1}^{3} \frac{\partial \hat{x}^p(\tau)}{\partial x^k(t)} \frac{\partial \hat{x}^p(\tau)}{\partial x^m(t)}$, $\rho_t$ is the mass density at time $t$, and $\mathscr{G}$ is a functional of the history of $C_t(\tau)$, $-\infty < \tau \leq t$, with a parametric dependence on $\rho_t$. The principle of material indifference requires that the functional $\mathscr{G}$ satisfies the following relation for an arbitary orthogonal matrix $Q$:

$$\underline{\underline{\mathscr{G}}}[\underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T;\ \rho_t] = \underline{\underline{Q}}\,\underline{\underline{\mathscr{G}}}[\underline{\underline{C}}_t(\tau);\ \rho_t]\,\underline{\underline{Q}}^T. \tag{7.7}$$

Consider now the following constitutive equation of a simple fluid: [4]

$$\underline{\underline{T}}(t) = -\hat{p}(\rho_t)\,\underline{\underline{1}} + \rho_t \int_{-\infty}^{t} \underline{\underline{C}}_t(\tau)\, \frac{\partial \hat{U}}{\partial \underline{\underline{C}}_t(\tau)}\, d\tau, \tag{7.8}$$

[4] The theory of such a fluid was proposed by Bernstein, Kearsley and Zapas [4]. Additional results on the same theory including the derivation of equation (7.8) appeared in a recent manuscript by Fong and Simmons [5].

where $\hat{p}$ is a scalar-valued function of the mass density $\rho_t$, and $\hat{U}$ is a scalar-valued function of one matrix argument $\mathbf{C}_t(\tau)$ and one scalar variable $t-\tau$, i.e., $U \equiv \hat{U}(\underline{\underline{C}}_t(\tau);\ t-\tau)$. Furthermore, $\hat{U}$ is required to satisfy the following condition for an arbitrary orthogonal $Q$:

$$\hat{U}(\underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T;\ t-\tau) = \hat{U}(\underline{\underline{C}}_t(\tau);\ t-\tau). \tag{7.9}$$

Differentiating eq (7.9) with respect to $\underline{\underline{C}}_t(\tau)$ and applying the formula No. 12 as listed in the last section, we obtain

$$\underline{\underline{Q}}^T\hat{\underline{\underline{U}}}_C(\underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T;\ t-\tau)\,\underline{\underline{Q}} = \hat{\underline{\underline{U}}}_C(\underline{\underline{C}}_t(\tau);\ t-\tau), \tag{7.10}$$

where $\hat{U}_C$ denotes the partial gradient of $\hat{U}$ with respect to $C_t(\tau)$. We are now ready to show that eq (7.8), indeed, satisfies the principle of material indifference. Using (7.6) and (7.8), we calculate the left-hand side of (7.7):

$$\underline{\underline{\mathscr{G}}}[\underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T;\ \rho_t] = -\hat{p}(\rho_t)\underline{\underline{1}} + \rho_t \int_{-\infty}^{t} \underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T\hat{\underline{\underline{U}}}_C(\underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T;\ t-\tau)\,d\tau.$$

Substituting (7.10) into the above equation, we obtain

$$\underline{\underline{\mathscr{G}}}[\underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T;\ \rho_t] = -\hat{p}(\rho_t)\underline{\underline{1}} + \rho_t \int_{-\infty}^{t} \underline{\underline{Q}}\,\underline{\underline{C}}_t(\tau)\,\underline{\underline{Q}}^T\underline{\underline{Q}}\,\underline{\underline{U}}_C(\underline{\underline{C}}_t(\tau);\ t-\tau)\,\underline{\underline{Q}}^T d\tau.$$

$$= \underline{\underline{Q}}\left\{-\hat{p}(\rho_t)\underline{\underline{1}} + \rho_t \int_{-\infty}^{t} \underline{\underline{C}}_t(\tau)\hat{\underline{\underline{U}}}_C(\underline{\underline{C}}_t(\tau);\ t-\tau)\,d\tau\right\}\underline{\underline{Q}}^T$$

$$= \underline{\underline{Q}}\,\underline{\underline{\mathscr{G}}}[\underline{\underline{C}}_t(\tau);\ \rho_t]\,\underline{\underline{Q}}^T = \text{R.H.S. of (7.7).} \qquad \textit{Q.E.D.}$$

It is clear from the above two examples that the advantage of adopting the matrix notation and applying formulas in "matrix calculus" lies mainly in the elegance in which higher-dimensional problems in continuum physics can be formulated. It is also clear that even though [illegible] of formulas was prepared for functions of one matrix argument, their applications can be easily extended to functions of several matrix arguments.

---

I wish to thank Dick Kraft, John A. Simmons, Seldon L. Stewart and Justin C. Walker for many helpful comments.

## 8. References

[1] Truesdell, C., and Noll, W., The Non-Linear Field Theories of Mechanics, Encyclopedia of Physics, **III/3**, Ed., S. Flügge (Springer-Verlag, 1965).
[2] Ericksen, J. L., Tensor Fields, An Appendix in Encyclopedia pf Physics, **III/1**, Ed., S. Flügge (Springer-Verlag, 1960).
[3] Newman, M., Integral Matrices (Academic Press (in press)).
[4] Bernstein, B., Kearsley, E. A., and Zapas, L. J., A Study of stress relaxation with finite strain, Transactions of Society of Rheology, **7**, 391 (1963).
[5] Fong, J. T., and Simmons, J. A., The scalar potential of the BKZ theory for a compressible fluid (to be published).

(Paper 75B3&4–349)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December, 1971

# Linear Transformations on Matrices*

**Marvin Marcus****

**(November 2, 1971)**

Let $K$ be a field and let $M_n(K)$ denote the vector space of all $n \times n$ matrices over $K$. Suppose $I(X)$ is an invariant defined on a subset $\mathfrak{A}$ of $M_n(K)$. This paper surveys certain results concerning the following problem. Describe the set $\mathcal{L}(I, \mathfrak{A})$ of all linear transformations $T: \mathfrak{A} \to \mathfrak{A}$ that hold the invariant $I$ fixed:

$$I(T(X)) = I(X), \qquad X \in \mathfrak{A}.$$

Key words: Matrices; invariants; determinant; generalized matrix function; rank.

## 1. Introduction

Let $K$ be a field and let $M_n(K)$ denote the vector space of all $n \times n$ matrices over $K$. Over the last 80 years, a great deal of effort has been devoted to the following question. Suppose $I(X)$ is an invariant defined on a subset $\mathfrak{A}$ of $M_n(K)$. Describe the set $\mathcal{L}(I, \mathfrak{A})$ of all linear transformations $T: \mathfrak{A} \to \mathfrak{A}$ that hold the invariant $I$ fixed:

$$I(T(X)) = I(X), \qquad X \in \mathfrak{A}. \tag{1}$$

Even in this generality, it is clear that $\mathcal{L}(I, \mathfrak{A})$ is a multiplicative semigroup with an identity. The invariant $I$ can be a scalar valued function, e.g., $I(X) = \det(X)$; or for that matter it can describe a property, e.g., $\mathfrak{A}$ can equal $M_n(C)$ and $I(X)$ can mean that $X$ is unitary, so that we are simply asking for the structure of all linear transformations $T$ that map the unitary group into itself.

Much of a beginning course in linear algebra is devoted to the study of one aspect of this question for certain choices of $I$; for example, if $I(X) = \rho(X)$, the rank of $X$, then it is well known that the three standard linear operations on the rows and columns of a matrix leave $\rho$ fixed and this fact permits us to compute $\rho(X)$ by reducing $X$ to some normal form. The similarity theory is another example of this problem. In this case take $I(X)$ to be the set of all elementary divisors of the characteristic matrix of $X$, and then the linear operators $T$ that we wish to study are precisely those for which $I(X) = I(T(X))$.

In the survey paper [18, 1962][1] some of the aspects of this general problem are discussed. But since the time that paper was written there have been a number of developments. The purpose of this paper is to describe some of these.

## 2. Survey of Results

The scalar invariants are functions $I$ for which $I(X)$ is either an element of $K$ (we will assume that char $K = 0$, so that integer-valued functions are included) or a $p$-tuple of elements of $K$. Prob-

---

*AMS Subject Classification:* Primary 15A72; Secondary 15A15 and 15A75.

* An invited paper. This research was supported by the U.S. Air Force Office of Scientific Research under Grant AFOSR 698–67.

** Present address: Department of Mathematics, University of California, Santa Barbara, California 93106.

[1] Figures in brackets indicate the literature references at the end of this paper.

ably the first three problems of this kind were considered by Frobenius [11, 1897]:

(i) $\mathfrak{A}=M_n(C)$, $\quad I(X)=\det(X)$;

(ii) $\mathfrak{A}$ is the space of real-symmetric (or odd order skew-symmetric) matrices and $I(X)=\det(X)$;

(iii) $\mathfrak{A}=\{X \mid \operatorname{tr}(X)=0\} \subset M_n(C)$ and $I(X)=\det(X)$.

Frobenius proved what one might expect, namely that $\mathscr{L}(\det, M_n(C))$ consists of linear transformations of the form

$$T(X)=UXV, \qquad X \epsilon M_n(C), \tag{2}$$

or

$$T(X)=UX^TV, \qquad X \epsilon M_n(C). \tag{3}$$

In problem (i) $\det(UV)=1$; in problem (ii) $U=\xi A$, $V=A^T$ where $\xi$ is an appropriately chosen constant and $\det A=1$; in (iii) $U=\xi A$, $V=A^{-1}$. I. Schur [34, 1925] extended and improved the result (i) as follows: Take $\delta(X)$ to be the $\binom{n}{k}^2$-tuple of $k^{th}$ order subdeterminants of $X$ in some order where $k \geqslant 3$ is fixed; then Schur proved that if $\delta(T(X))=S(\delta(X))$ for $S$ a fixed nonsingular matrix, then $T$ has one of the two forms (2) or (3) (without the restriction $\det(UV)=1$). Dieudonné [8, 1949] showed that if $T$ is a semi-linear transformation of $M_n(K)$ onto itself which holds the cone $\det(X)=0$ invariant, then $T$ is of the form

$$T(X)=U(\sigma(x_{ij}))V, \qquad X \epsilon M_n(K),$$

or

$$T(X)=U(\sigma(x_{ij}))^TV, \qquad X \epsilon M_n(K),$$

where $\sigma$ is an automorphism of the field. In the paper [9, 1957] Dynkin states that the Frobenius theorems can be obtained using some results on the structure of maximal subgroups of the classical groups.

In an old result, Pòlya [32] restricted $T$ to be a linear transformation which affixes in a prescribed way $+$ and $-$ signs to the elements of $X$, and asked whether such a $T$ exists which satisfies $\operatorname{per}(T(X))=\det(X)$ for $n>2$. Pòlya answered the question negatively and many years later in [23, 1961] the question was answered negatively for arbitrary linear transformations $T$. Along the lines of the Frobenius and Schur results, the structure of $\mathscr{L}(E_r, M_n(C))$ is determined where $E_r(X)$ is the $r$th elementary symmetric function of the eigenvalues of $X$, i.e., the sum of all $r$-square principle subdeterminants of $X$. It was proved in [28, 1959] that for $4 \leqslant r < n$ any $T \epsilon \mathscr{L}(E_r, M_n(C))$ is of the form

$$T(X)=UXV, \qquad X \epsilon M_n(C),$$

or

$$T(X)=UX^TV, \qquad X \epsilon M_n(C),$$

where $UV=e^{i\varphi}I_n$ and $r\varphi \equiv 0\,(2\pi)$. Just recently Beasley [1, 1970] completed the argument by showing that for $r=3$, precisely the same result holds. E. P. Botta, in several papers considers the choice $I(X)=d(X)$ where $d$ is a generalized matrix function in the sense of Schur, i.e.,

$$d(X)=\sum_{\sigma \epsilon \Pi} \lambda(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)},$$

where $\lambda$ is a nonzero function defined on a subgroup $H$ of $S_n$. In [3, 1967] Botta determined the structure of $\mathscr{L}(d, M_n(K))$ when $H$ is a transitive cyclic subgroup of $S_n$. In [4, 1968] $H$ is taken to be a doubly transitive or regular proper subgroup of $S_n$ and $\lambda$ is a character of $H$ of degree 1. In [5, 1967] Botta reproves an earlier result of Marcus and May [22, 1962] showing that $\mathscr{L}$ (per, $M_n(K)$) consists of precisely those $T$ of the form

$$T(X) = DPXQL, \qquad X \in M_n(K),$$

or

$$T(X) = DPX^{\tau}QL, \qquad X \in M_n(K),$$

where $P$ and $Q$ are permutation matrices, $D$ and $L$ are diagonal matrices and per $(DL) = 1$.

Many of the results concerning the structure of $\mathscr{L}(I, \mathfrak{A})$ can be reduced to the problem of determining linear maps on $M_n(K)$ which map the set of rank 1 matrices into itself. W. L. Chow [6, 1949], L. K. Hua [12, 1951] and Jacobson and Rickart [13, 1950] considered $1-1$ onto maps $T$ of $M_n(K)$ which have the property that both $T$ and $T^{-1}$ preserve coherence. In the present context, this amounts to assuming that $T$ and $T^{-1}$ both have the property that whenever $X$ and $Y$ differ by a matrix of rank 1, then $T(X)$ and $T(Y)$ differ by a matrix of rank 1. For linear maps this means $\rho(T(X)) = \rho(X)$ for all $X$. In [26, 1959] Marcus and Moyls proved that if $T: M_{m,n}(K) \to M_{m,n}(K)$ ($M_{m,n}(K)$ is the space of all $m \times n$ matrices over $K$) is linear and $\rho(T(X)) = 1$ whenever $\rho(X) = 1$, then $T$ has the form

$$T(X) = UXV, \qquad X \in M_{m,n}(C),$$

or

$$T(X) = UX^{\tau}V, \qquad X \in M_{m,n}(C),$$

where $U$ and $V$ are fixed nonsingular matrices in $M_m(K)$ and $M_n(K)$ respectively. This result is fairly easy to apply because it does not require the a priori existence of $T^{-1}$. R. Westwick in [36, 1967] extended the result in [26, 1959] to linear maps on the space of $n$-contravariant tensors which hold the nonzero decomposable elements set-wise fixed. In another paper [35, 1964] Westwick, using techniques in [6, 1949] determined the structure of linear maps on the space $\wedge^m V$ of skew-symmetric tensors into itself which hold the set of nonzero decomposable elements set-wise fixed. In a thesis at the University of British Columbia [7, 1967] L. Cummings proved that if $T$ maps the symmetric power $V^{(m)}$ into itself and holds the set of non-zero decomposable elements set-wise fixed then $T$ is induced by a linear map of V. Cummings' result requires that the underlying field be algebraically closed of characteristic either 0 or exceeding $m$. In another thesis [14, 1971] M. H. Lim reconsiders this problem and relaxes the conditions on the field. Beasley [2, 1970] considered the problem of determining all linear transformations $T: M_n(K) \to M_n(K)$, $K$ algebraically closed, which hold the set of rank $k$ matrices set-wise fixed. Beasley required additional hypotheses on $T$ in order to prove that $T$ has the form (2) or (3). Djoković [10, 1969] proved that if $T$ maps the set of rank $k$ matrices into itself and is nonsingular, then in fact $T$ maps the set of rank 1 matrices into itself and the result in [26, 1959] applies. Much earlier [30, 1941] Morita proved that if $T$ maps the set of rank 1 matrices into itself and maps the set of rank 2 matrices into the set of matrices of rank at least 2, then $T$ has the form (2) or (3). He then used this to prove a result of Schur to the effect that if $I(X) = \alpha_1(X)$ is the Hilbert norm of $X$, i.e., the square root of the largest eigenvalue of $X^*X$, and $T \in \mathscr{L}(I, M_{m,n}(C))$, then $T$ has the form (2) or (3) in which $U$ and $V$ are unitary. In a later paper [31, 1944], Morita shows that if $\mathfrak{A}$ is the set of n-square complex skew-symmetric matrices and $I(X)$ is again the Hilbert norm of $X$, then for $n \neq 4$ and $T \in \mathscr{L}(I, \mathfrak{A})$,

$$T(X) = U^{\tau}XU, \qquad X \in M_n(C),$$

where $U$ is a fixed unitary matrix; or if $n=4$, then $T(X)$ can also have the alternative form

$$T(X)=U^{\tau}X^{+}U, \qquad X\epsilon M_n(C), \tag{4}$$

where $X^{+}$ is the matrix obtained from $X$ by interchanging the (1, 4) and the (2, 3) entry. A result similar to this was obtained by Westwick in his thesis [35, 1964]. In [29, 1960] Marcus and Westwick proved a theorem somewhat along the lines of the Morita theorem as follows. Let $k$ be a fixed integer satisfying $4 \leq 2k \leq n$. Let $\mathfrak{A}$ be the set of skew-symmetric matrices over the field $R$ of real numbers and let $T\epsilon\mathscr{L}(E_{2k}, \mathfrak{A})$. If $n \geq 5$, then there exists a real matrix $P$ such that

$$T(X)=\alpha PXP^{\tau}, \qquad X\epsilon\mathfrak{A}, \tag{5}$$

where $\alpha PP^{\tau}=I_n$ is $2k<n$ and $\alpha PP^{\tau}$ is unimodular if $2k=n$. If $2k=n=4$, then either $T$ has the form (5) or

$$T(X)=\alpha P\begin{bmatrix} 0 & x_{34} & x_{24} & x_{23} \\ -x_{12} & 0 & x_{23} & x_{24} \\ -x_{13} & -x_{23} & 0 & x_{34} \\ -x_{14} & -x_{24} & -x_{34} & 0 \end{bmatrix}P^{\tau},$$

where

$$X=\begin{bmatrix} 0 & x_{12} & x_{13} & x_{14} \\ -x_{12} & 0 & x_{23} & x_{24} \\ -x_{13} & -x_{23} & 0 & x_{34} \\ -x_{14} & -x_{24} & -x_{34} & 0 \end{bmatrix},$$

and $\alpha PP^{\tau}$ is unimodular. Later on Marcus and Minc [24, 1962] proved that if $T\epsilon\mathscr{L}(E_r', M_{m,n}(C))$ where $1<r\leq n$ and $E_r'(X)$ is the value of $E_r$ at the squares of the singular values of $X$, i.e., $E_r'(X)$ is just the value of $E_r$ at the eigenvalues of $X^{*}X$, then $T$ has the form (2) if $m\neq n$ and either (2) or (3) if $m=n$, where $U\epsilon M_n(C)$ and $V\epsilon M_n(C)$ are unitary.

In [17, 1959] it is proved that if $T$ is a linear map of $M_n(C)$ into itself such that $T(X)$ is unitary whenever $X$ is unitary, then $T$ is of the form (2) or (3) where $U$ and $V$ are unitary. B. Russo [33, 1969] recently used this result to prove the following interesting theorem. If $I(X)$ is the sum of the singular values of $X$ and if $T$ maps the identity matrix into itself, then $T\epsilon\mathscr{L}(I, M_n(C))$ has the form (2) or (3), where $U$ and $V$ are unitary. Marcus and Gordon [20, 1970] recently proved the following result. Let $f(t)=f(t_1, \ldots, t_n)$ be a continuous, real-valued function defined for all $t_j \geq 0$, $1 \leq j \leq n$, and for $X\epsilon M_{m,n}(C)$, let

$$I(X)=f(\alpha_1(X), \alpha_2(X), \ldots, \alpha_n(X))$$

where $\alpha_1(X) \geq \alpha_2(X) \geq \ldots \geq \alpha_n(X)$ are the singular values of $X$. If $f(t_1, \ldots, t_n)$ is concave, symmetric, strictly increasing in each $t_j$, and $f(0)=0$, then $T\epsilon\mathscr{L}(I, M_{m,n}(C))$ has the form (2) if $m\neq n$ and either (2) or (3) if $m=n$ where $U\epsilon M_m(C)$ and $V\epsilon M_n(C)$ are unitary. By specializing $f$ to

$$f(t)=\sum_{j=1}^{n} t_j^{\sigma}$$

where $0<\sigma\leq 1$, the above theorem reduces to the following result. If $T: M_{m,n}(C) \to M_{m,n}(C)$ satisfies

$$\sum_{j=1}^{n} \alpha_j(T(X))^{\sigma}=\sum_{j=1}^{n} \alpha_j(X)^{\sigma}$$

for all $X \in M_{m,n}(C)$, then (2) or (3) holds with unitary $U$ and $V$. In the paper [27, 1970] the following result is proven using representation theory techniques. Let $f(t) = f(t_1, \ldots, t_n)$ satisfy the conditions

(i) $f(t) = 0$ if and only if $t = 0$;

(ii) $f$ is positively homogeneous of degree $\rho \neq 0$; i.e., $f(ct) = c^\rho f(t)$, all $c \geq 0$, $t \geq 0$ (i.e., $t_j \geq 0, j = 1, \ldots, n$).

If $I(X) = f(\alpha_1(X), \ldots, \alpha_n(X))$ as before, then $\mathscr{L}(I, M_{m,n}(C))$ is a subgroup of the group of $mn \times mn$ unitary matrices $U(mn, C)$ where we associate each $T \in \mathscr{L}(I, M_{m,n}(C))$ with its matrix representation with respect to the lexicographically ordered basis $\{E_{st} = (\delta_{is}\delta_{tj}), i, j = 1, \ldots, n\}$.

## 3. Current Work and Some Questions

M. J. S. Lim, in work closely related to that of Marcus and Westwick [29, 1960], has recently published [15, 1970] the following result. Let $T$ map the space of skew-symmetric matrices over an algebraically closed field $K$ into itself. Assume that $T$ maps the set of rank 4 matrices into itself. Then for $n \neq 4$, $T$ is of the form

$$T(X) = \alpha P X P^T$$

or

$$T(X) = \alpha P X^T P^T.$$

In case $n = 4$, $T$ is one of the above forms, or else

$$T(X) = \alpha P X^+ P^T$$

where $X^+$ is defined in (4).

Just recently Marcus and Holmes [21, 1971] have proved the following results. Let $X \in M_n(C)$. For any subgroup $H$ of the symmetric group $S_m$ of degree $m$ and character $\chi$ of degree 1 on $H$, let $K(X) : P \to P$ be the induced transformation [19, 1967] on the symmetry class of tensors $(P, \nu)$ associated with $H$ and $\chi$. Define $I(X) = tr\, K(X), X \in M_n(C)$.

(i) Let $m \leq n$ or $\chi \equiv 1$. $\mathscr{L}(I, M_n(C))$ is a group if and only if $H \neq \{e\}$.

(ii) Let $H = S_m$, $\chi \equiv 1$ and $\mathfrak{A} \subset M_n(C)$ an algebra with the property that $\mathfrak{A}^* = \{X^* : X \in \mathfrak{A}\} = \mathfrak{A}$, when $X^*$ denotes the conjugate transpose of $X$. Then $\mathscr{L}(I, \mathfrak{A})$ is a group.

(iii) In (i) take $H = S_m$, $m \geq 3$ and $\chi \equiv 1$. If $\mathscr{L}_1(I, M_n(C))$ denotes the subgroup of $\mathscr{L}(I, M_n(C))$ of those $T : M_n(C) \to M_n(C)$ satisfying $T(I_n) = \xi I_n$, then for any $T \in \mathscr{L}_1(I, M_n(C))$ there exists a fixed nonsingular matrix $P \in M_n(C)$ such that

$$T(X) = \xi P^{-1} X P, \qquad X \in M_n(C), \tag{6}$$

or

$$T(X) = \xi P^{-1} X^T P, \qquad X \in M_n(C). \tag{7}$$

In this case $tr\, K(X)$ is the completely symmetric function of the eigenvalues of $X$, denoted here by $h_m(X)$. Thus this result states that if $T(I_n) = \xi I_n$ and $h_m(T(X)) = h_m(X)$ for all $X \in M_n(C)$ then $T$ has the form (6) or (7).

(iv) In (i) take $A_m \subset S_m$ to be the alternating group, $m \geq 3$ and $\chi \equiv 1$. Then the group $\mathscr{L}_1(I, M_n(C))$ consists precisely of those linear transformations $T$ of the form (6) or (7).

There are a number of questions which remain unanswered. For example, a more direct proof of the result in [27, 1970] might be based on the following.

CONJECTURE 1: *Let* T *be an* mn*-square complex matrix, and assume that for arbitrary unitary matrices* $U \in M_n(C)$, $V \in M_n(C)$ *the matrix* $(U \otimes V)T$ *has eigenvalues of modulus 1. (The matrix* $U \otimes V$ *is the usual Kronecker product of* U *and* V.*) Then* T *is unitary.*

CONJECTURE 2: *If* $T: M_n(C) \to M_n(C)$ *is a linear map and if* $h_m(T(X)) = h_m(X)$, $X \in M_n(C)$ *(recall that* $h_m(X)$ *is the completely symmetric function of the eigenvalues of* X*), then in fact* $T(I_n) = \xi I_n$ *and hence from* [21, 1971] T *has the form (6) or (7).*

CONJECTURE 3: *Let* $P_m(X)$ *denote the* mth *induced power matrix of* X [25] *and suppose that* $T: M_n(C) \to M_n(C)$ *satisfies* $P_m(T(X)) = S(P_m(X))$, $X \in M_n(C)$ *where* $S: M_N(C) \to M_N(C)$ *is a fixed non-singular linear map,* $N = \binom{m+m-1}{m}$. *Then* T *has the form (2) or (3).*

CONJECTURE 4: *Suppose* K(X) *is an invariant matrix* [16, Chapter X] *defined by means of a Young tableau. If* $T: M_n(C) \to M_n(C)$ *and* $\operatorname{tr} K(T(X)) = \operatorname{tr} K(X)$ *for all* $X \in M_n(C)$, *then* T *must have the form (2) or (3) with some appropriate conditions on the* U *and* V. *Of course, the result in* [28, 1959] *and* [21, 1971] *are special cases of this.*

*As a possible extension of Schur's theorem* [30, 1941] *consider*

CONJECTURE 5: *Let* $T: M_n(C) \to M_n(C)$ *and* $h_m(T(X)^*T(X)) = h_m(X^*X)$ *(see Conjecture 2 in which* $h_m$ *is defined), then* T *has the form (2) or (3).*

*As a variant of the result in* [17, 1959], *let* G *be any of the following classical groups: the real orthogonal group, the rotation group, the symplectic group.*

CONJECTURE 6: *Let* $T: M_n(R) \to M_n(R)$ *map* G *into itself. Then* T *must have the form (2) or (3) in which* U *and* V *belong to* G. *(In the case of the rotation group,* U *and* V *could be simply real orthogonal with* $\det(UV) = 1$.)

CONJECTURE 7: *Suppose* T *is a mapping of the space of 2-contravariant tensors into itself; and suppose moreover that for each decomposable element* $x \otimes y$ *we have* $\|T(x \otimes y)\| = \|x \otimes y\|$ *(Euclidean norm). Then* T *is unitary. This can be restated in terms of linear maps* $T: M_n(C) \to M_n(C)$. *Thus suppose for* $\rho(X) = 1$, $\|T(X)\| = \|X\|$ *where* $\|X\|$ *is just the Euclidean norm. Show that* T *is unitary.*

## 4. References


[1] Beasley, L., Linear transformations on matrices: The invariance of the third elementary symmetric function, Canad. J. Math. **22** (1970), 746–752.

[2] Beasley, L., Linear transformations on matrices: The invariance of rank $k$ matrices, Linear Algebra and Appl. **3** (1970), 407–427.

[3] Botta, E., Linear transformations on matrices: The invariance of a class of general matrix functions, Canad. J. Math. **19** (1967), 281–290.

[4] Botta, E., Linear transformations on matrices: The invariance of a class of general matrix functions II, Canad. J. Math. **20** (1968), 739–748.

[5] Botta, E., Linear transformations that preserve the permanent, Proc. Amer. Math. Soc. **18** (1967), 566–569.

[6] Chow, W. L., On the geometry of algebraic homogeneous spaces, Ann. Math. **50** (1949), 32–67.

[7] Cummings, L., Linear transformations on symmetric tensor spaces which preserve rank 1 (Ph.D. thesis), University of British Columbia (1967).

[8] Dieudonné, J., Sur une généralisation du groupe orthogonal à quatre variables, Arch. Math. **1** (1949), 282–287.

[9] Dynkin, E. B., The maximal subgroups of the classical groups, Amer. Math. Soc. Translations Series 2, **6** (1959).

[10] Djoković, D., Linear transformations of tensor products preserving a fixed rank, Pacific J. Math. **30** (1969), 411–414.

[11] Frobenius, G., Über die Darstellung der endlichen Gruppen durch linear Substitutionen: I, Sitzungsberichte der Preuss. Acad. Wiss. zu Berlin (1897), 994–1015.

[12] Hua, L. K., A theorem on matrices over a field and its applications, J. Chinese Math. Soc. (N.S.) **1** (1951), 110–163.

[13] Jacobson, N., and Rickart, C., Jordan homomorphisms of rings, Trans. Amer. Math. Soc. **69** (1950), 479–502.

[14] Lim, M. H., (Ph.D. thesis, in preparation) University of British Columbia (1971).

[15] Lim, M. J. S., Rank preservers of skew-symmetric matrices, Pacific J. Math. **35** (1970), 169–174.

[16] Littlewood, D., The Theory of Group Characters and Matrix Representations of Groups, (Oxford University Press, 1940).

[17] Marcus, M., All linear operators leaving the unitary group invariant, Duke Math. J. **26** (1959), 155–163.

[18] Marcus, M., Linear operations on matrices, Amer. Math. Monthly **69** (1962), 837–847.

[19] Marcus, M., A theorem on rank with applications to mappings on symmetry classes of tensors, Bull. Amer. Math. Soc. **73** (1967), 675–677.

[20] Marcus, M., and Gordon, W., A generalization of the unitary group, Linear Algebra and Appl. **3** (1970), 225–247.

[21] Marcus, M., and Holmes, J., Groups of linear operators defined by group characters, in press.

[22] Marcus, M., and May, F., The permanent function, Canad. J. Math. **14** (1962), 177–189.

[23] Marcus, M., and Minc, H., On the relation between the determinant and the permanent, Ill. J. Math. **5** (1961), 376–381.

[24] Marcus, M., and Minc, H., The invariance of symmetric functions of singular values, Pacific J. Math. **12** (1962), 327–332.

[25] Marcus, M., and Minc, H., A Survey of Matrix Theory and Matrix Inequalities (Prindle, Weber and Schmidt, 1964).

[26] Marcus, M., and Moyls, B., Transformations on tensor product spaces, Pacific J. Math. **9** (1959), 1215–1221.

[27] Marcus, M., and Newman, M., Some results on unitary matrix groups, Linear Algebra and Appl. **3** (1970), 173–178.

[28] Marcus, M., and Purves, R., Linear transformations on algebras of matrices II: The invariance of the elementary symmetric functions, Canad. J. Math. **11** (1959), 383–396.

[29] Marcus, M., and Westwick, R., Linear maps on skew-symmetric matrices: The invariance of elementary symmetric functions, Pacific J. Math. **10** (1960), 917–924.

[30] Morita, K., Analytical characterization of displacements in general Poincaré space, Proc. Imperial Acad. **17** (1941), 489–494.

[31] Morita, K., Schwarz's lemma in a homogeneous space of higher dimension, Japanese J. Math. **19** (1944), 45–56.

[32] Pòlya, G., Aufgabe 424, Arch. Math. u. Phys. **203**, 271.

[33] Russo, B., Trace preserving mappings of matrix algebras. Duke Math. J. **36** (1969), 297–300.

[34] Schur, I., Einige Bemerkungen zur Determinanten Theorie, Sitzungsberichte der Preuss. Akad. Wiss. zu Berlin **25** (1925), 454–463.

[35] Westwick, R., Linear transformations on Grassmann spaces, Pacific J. Math. **14** (1964), 1123–1127.

[36] Westwick, R., Transformations on tensor spaces, Pacific J. Math. **23** (1967), 613–620.


(Paper 75B3 & 4–350)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# Eigenvalues of Sums of Hermitian Matrices. III.*

**Robert C. Thompson** and Linda Freede Garbanati****

November 2, 1971

Two classes of nonlinear inequalities for the eigenvalues of sums of Hermitian matrices are obtained. These nonlinear inequalities are shown to follow from linear inequalities established in parts I and II of this series. A new inequality for the singular values of matrix products is also obtained.

Keywords: Eigenvalues; linear inequalities; singular values.

## 1. Introduction.

Let $C=A+B$ where $C, A, B$ are Hermitian matrices. In recent years a number of inequalities have been established linking the eigenvalues of $C, A, B$. For the most part these inequalities are linear, but a number of nonlinear inequalities are also known which involve convex or concave functions of the eigenvalues. Generally speaking, the methods used to derive the linear inequalities work equally well in deriving these nonlinear inequalities. These methods are based on the extremal properties of eigenvalues or on induction and the Cauchy interlacing inequalities. A defect of the use of these methods when deriving nonlinear inequalities is that they do not clarify whether the nonlinear inequalities are consequences of the linear ones. There is, however, a third method which derives nonlinear inequalities from linear ones. It is the purpose of this paper to use this third method to obtain several new families of nonlinear inequalities for the eigenvalues of matrix sums. It will turn out that our new inequalities include as a special case a generalized and sharpened version of an inequality recently proved by Marcus [2].[1] We shall also derive a previously unnoticed inequality for the singular values of a matrix product.

## 2. The Main Results for Sums

THEOREM 1: *Let integers* $i_1, \ldots, i_m, j_1, \ldots, j_m$ *satisfy*

$$1 \leq i_1 < \ldots < i_m \leq n, \qquad 1 \leq j_1 < \ldots < j_m \leq n, \qquad i_m + j_m - m \leq n.$$

*Set*

$$k_s = i_s + j_s - s \qquad \textit{for} \qquad s = 1, \ldots, m.$$

*Let* $0 \leq \theta \leq 1$, $\theta + \varphi = 1$ *and let* $A, B, C = \theta A + \varphi B$ *be n-square Hermitian matrices with eigenvalues*

$$\alpha_1 \geq \ldots \geq \alpha_n, \beta_1 \geq \ldots \geq \beta_n, \gamma_1 \geq \ldots \geq \gamma_n, \tag{1}$$

---

*AMS Subject Classification:* Primary: 15A42; Secondary: 15A39.

*An invited paper. The preparation of this paper was supported in part by the U.S. Air Force Office of Scientific Research, under Grant 698–67.

**Present address: Department of Mathematics, University of California, Santa Barbara, California 93106.

[1] Figures in brackets indicate the literature references at the end of this paper.

*respectively. Let* $f(x_1, \ldots, x_m)$ *be a real-valued function, convex, symmetric, and nondecreasing in each* $x_t$ *for* $(x_1, \ldots, x_m)$ *lying in the* m*-dimensional hypercube* $\mathcal{I} \times \ldots \times \mathcal{I}$, *where* $\mathcal{I}$ *is the interval*

$$\mathcal{I} = [min\ (\alpha_n, \beta_n), max\ (\alpha_1, \beta_1)].$$

*Then*

$$f(\gamma_{k_1} \ldots, \gamma_{k_1}) \leq \theta f(\alpha_{i_1}, \ldots, \alpha_{i_m}) + \varphi f(\beta_{j_1}, \ldots \ldots \beta_{j_m}). \tag{2}$$

PROOF: The inequality

$$\sum_{s=1}^{m} \gamma_{k_s} \leq \theta \sum_{s=1}^{m} \alpha_{i_s} + \varphi \sum_{s=1}^{m} \beta_{j_s} \tag{3}$$

is known [7, Theorem 1]. Applying this for $r$ in place of $m$, we have

$$\sum_{s=1}^{r} \gamma_{k_s} \leq \theta \sum_{s=1}^{r} \alpha_{i_s} + \varphi \sum_{s=1}^{r} \beta_{j_s}, \qquad r = 1, \ldots, m.$$

If we let

$$c_s = \gamma_{k_s}, d_s = \theta \alpha_{i_s} + \varphi \beta_{j_s}, \qquad s = 1, \ldots, m,$$

then $c_1 \geq \ldots \geq c_m$, $d_1 \geq \ldots \geq d_m$, and $c_1 + \ldots + c_r \leq d_1 + \ldots + d_r$ for $1 \leq r \leq m$. These conditions are known [5, Theorem 15] to imply that $f(c_1, \ldots, c_m) \leq f(d_1, \ldots, d_m)$. From this (2) is immediate.

This proof uses only standard techniques. Rather more interesting and surprising is that these same techniques can be used to establish the inequalities of the next theorem.

THEOREM 2: *Let* $C = \theta A + \varphi B$ *and* f *be as in Theorem 1. Let* $Z_1, \ldots, Z_{n-m}, W_1, \ldots, W_{n-m}$ *be integers satisfying*

$$Z_1 \leq \ldots \leq Z_{n-m} \leq m, \qquad W_1 \leq \ldots \leq W_{n-m} \leq m, \qquad Z_1 + W_1 \geq m.$$

*Let* $\delta_x(y)$ *be a jump function;* $\delta_x(y) = 1$ *if* $y > x$ *and* $\delta_x(y) = 0$ *if* $y \leq x$. *Let*

$$I_s = s + \delta_{Z_1}(s) + \ldots + \delta_{Z_{n-m}}(s),$$

$$J_s = s + \delta_{W_1}(s) + \ldots + \delta_{W_{n-m}}(s),$$

$$K_s = s + \delta_{Z_1+W_1-m}(s) + \ldots + \delta_{Z_{n-m}+W_{n-m}-m}(s), \qquad s = 1, \ldots, m.$$

*Then*

$$f(\gamma_{K1}, \ldots, \gamma_{Km}) \leq \theta f(\alpha_{I1}, \ldots, \alpha_{Im}) + \phi f(\beta_{J1}, \ldots, \beta_{Jm}). \tag{4}$$

PROOF: The following inequality is known:

$$\sum_{s=1}^{m} \gamma_{K_s} \leq \theta \sum_{s=1}^{m} \alpha_{I_s} + \phi \sum_{s=1}^{m} \beta_{J_s}. \tag{5}$$

This inequality may be obtained from Theorem 4 of [8] by using $\delta_x^*(y) = \delta_{x-1}(y)$. In order to prove (4) we need to show that

$$\sum_{s=1}^{r} \gamma_{K_s} \leq \theta \sum_{s=1}^{r} \alpha_{I_s} + \phi \sum_{s=1}^{r} \beta_{J_s}. \tag{6}$$

holds for each $r=1, \ldots, m$. For if this is so then the proof of Theorem 1 applies here also. Define nonnegative integers $u$ and $v$ by

$$Z_1 \leq \ldots \leq Z_u \leq r < Z_{u+1} \leq \ldots \leq Z_{n-m},$$

$$W_1 \leq \ldots \leq W_v \leq r < W_{v+1} \leq \ldots \leq W_{n-m}.$$

Without loss of generality let $u \leq v$. Define $\xi_1 \leq \ldots \leq \xi_{n-r}$, $\omega_1 \leq \ldots \leq \omega_{n-r}$ by

$$\begin{gathered} \xi_1 = Z_1 - (m-r), \ldots, \xi_u = Z_u - (m-r), \xi_{u+1} = r, \ldots, \xi_{n-r} = r, \\ \omega_1 = W_1, \ldots, \omega_v = W_v, \omega_{v+1} = r, \ldots, \omega_{n-r} = r. \end{gathered} \tag{7}$$

Then also $\xi_1 + \omega_1 \geq r$. Set

$$P_s = s + \delta_{\xi_1}(s) + \ldots + \delta_{\xi_{n-r}}(s),$$

$$Q_s = s + \delta_{\omega_1}(s) + \ldots + \delta_{\omega_{n-r}}(s),$$

$$R_s = s + \delta_{\xi_1+\omega_1-r}(s) + \ldots + \delta_{\xi_{n-r}+\omega_{n-r}-r}(s), \qquad s = 1, \ldots, r.$$

Writing down the inequality (5) with $r$ in place of $m$ we have

$$\sum_{s=1}^{r} \gamma_{R_s} \leq \theta \sum_{s=1}^{r} \alpha_{P_s} + \phi \sum_{s=1}^{r} \beta_{Q_s}. \tag{8}$$

Now for $s \leq r$ we have

$$P_s = s + \sum_{\rho=1}^{n-r} \delta_{\xi_\rho}(s) = s + \sum_{\rho=1}^{u} \delta_{\xi_\rho}(s) \geq s + \sum_{\rho=1}^{u} \delta_{Z_\rho}(s) = I_s,$$

because

$$\delta_{\xi_\rho}(s) = 0 \text{ for } s \leq r \quad \text{and} \quad \rho > u;$$

$$\delta_{\xi_\rho}(s) = \delta_{Z_\rho - (m-r)}(s) \geq \delta_{Z_\rho}(s) \quad \text{for} \quad \rho \leq u.$$

For $s \leq r$ we have

$$Q_s = s + \sum_{\rho=1}^{n-r} \delta_{\omega_\rho}(s) = s + \sum_{\rho=1}^{v} \delta_{\omega_\rho}(s) = s + \sum_{\rho=1}^{v} \delta_{W_\rho}(s) = s + \sum_{\rho=1}^{n-m} \delta_{W_\rho}(s) = J_s,$$

since

$$\delta_{\omega_\rho}(s) = 0 = \delta_{W_\rho}(s) \quad \text{for} \quad \rho > v.$$

Furthermore, for $s \leq r$ (using $Z_\rho + W_\rho - m = \xi_\rho + \omega_\rho - r$ for $\rho \leq u$ *and* $Z_\rho - m + W_\rho \leq \omega_\rho$ for all $\rho$) we have

$$K_s \geq s + \sum_{\rho=1}^{u} \delta_{Z_\rho + W_\rho - m}(s) + \sum_{\rho=u+1}^{v} \delta_{Z_\rho + W_\rho - m}(s) \geq s + \sum_{\rho=1}^{u} \delta_{\xi_\rho + \omega_\rho - r}(s) + \sum_{\rho=u+1}^{v} \delta_{\omega_\rho}(s)$$

$$= s + \sum_{\rho=1}^{u} \delta_{\xi_\rho + \omega_\rho - r}(s) + \sum_{\rho=u+1}^{n-r} \delta_{\xi_\rho + \omega_\rho - r}(s) = R_s.$$

Thus, for $s=1, \ldots, r$, we have $P_s \geq I_s$, $Q_s = J_s$, $K_s \geq R_s$. From this it follows that (6) is a consequence of (8). This completes the proof.

REMARK 1: In fact we have proved that if $f(x_1, \ldots, x_r)$ is convex, symmetric, and nondecreasing then

$$f(\gamma_{K_1}, \ldots, \gamma_{K_r}) \leq \theta f(\alpha_{I_1}, \ldots, \alpha_{I_r}) + \varphi f(\beta_{J_1}, \ldots, \beta_{J_r}) \qquad \text{for} \qquad r \leq m. \tag{9}$$

That the inequality (9) for $r < m$ is valid is a new result even when $f$ is linear.

As a consequence of theorems 1 and 2, we obtain theorems 3 and 4.

THEOREM 3: *Let* A, B, C *be as in theorem 1. Let* $g(x_1, \ldots, x_m)$ *be a real-valued function, symmetric, convex, and nonincreasing in each variable when* $(x_1, \ldots, x_m) \in \mathfrak{I} \times \ldots \times \mathfrak{I}$. *Let* $i'_1, \ldots, i'_m, j'_1, \ldots, j'_m, k'_1, \ldots, k'_m$ *be integers with*

$$1 \leq i'_1 < \ldots < i'_m \leq n, \qquad 1 \leq j'_1 < \ldots < j'_m \leq n, \qquad i'_1 + j'_1 > n+1-m,$$

$$k'_s = i'_s + j'_s - s + m - n, \qquad 1 \leq s \leq m.$$

*Then*

$$g(\gamma_{k'_1}, \ldots, {}_{k'_m}) \leq \theta g(\alpha_{i'_1}, \ldots, \alpha_{i'_m}) + \varphi g(\beta_{j'_1}, \ldots, \beta_{j'_m}).$$

THEOREM 4: *Let* A, B, C, g *be as in theorem 3. Let integers* $Z'_1, \ldots, Z'_{n-m}, W'_1, \ldots, W'_{n-m}$ satisfy

$$0 \leq Z'_1 \leq \ldots \leq Z'_{n-m}, \qquad 0 \leq W'_1 \leq \ldots \leq W'_{n-m}, \qquad Z'_{n-m} + W'_{n-m} \leq m.$$

*Let*

$$I'_s = s + \sum_{\rho=1}^{n-m} \delta_{Z'_\rho}(s), \qquad J'_s = s + \sum_{\rho=1}^{n-m} \delta_{W'_\rho}(s),$$

$$K'_s = s + \sum_{\rho=1}^{n-m} \delta_{Z'_\rho + W'_\rho}(s), \qquad s = 1, \ldots, m.$$

*Then*

$$g(\gamma_{K'_1}, \ldots, \gamma_{K'_m}) \leq \theta g(\alpha_{I'_1}, \ldots, \alpha_{I'_m}) + \varphi g(\beta_{J'_1}, \ldots, \beta_{J'_m}).$$

PROOF: Let $f(x_1, \ldots, x_m) = g(-x_1, \ldots, -x_m)$ and apply theorems 1 and 2 to $-C = \theta(-A) + \varphi(-B)$. This yields theorem 3 in a straightforward manner. To obtain theorem 4 take $Z_s = m - Z'_{m+1-s}$ and use $1 - \delta_z(m+1-s) = \delta_{m-z}(s)$.

From theorems 1–4 we deduce

$$F(\gamma_{k_1}, \ldots, \gamma_{k_m}) \geq \theta F(\alpha_{i_1}, \ldots, \alpha_{i_m}) + \varphi F(\beta_{j_1}, \ldots, \beta_{j_m})$$

$$F(\gamma_{K_1}, \ldots, \gamma_{K_m}) \geq \theta F(\alpha_{I_1}, \ldots, \alpha_{I_m}) + \varphi F(\beta_{J_1}, \ldots, \beta_{J_m})$$

$$G(\gamma_{k_{i'}}, \ldots, \gamma_{k_{m'}}) \geq \theta G(\alpha_{i_{i'}}, \ldots, \alpha_{i_{m'}}) + \varphi G(\beta_{j_{i'}}, \ldots, \beta_{j_{m'}})$$

$$G(\gamma_{K_{i'}}, \ldots, \gamma_{K_{m'}}) \geq \theta G(\alpha_{I_{i'}}, \ldots, \alpha_{I_{m'}}) + \varphi G(\beta_{J_{i'}}, \ldots, \beta_{J_{m'}})$$

where $F$ is symmetric, concave, and nonincreasing on $\mathscr{I}\times\ldots\times\mathscr{I}$, and $G$ is symmetric, concave, and nondecreasing on $\mathscr{I}\times\ldots\times\mathscr{I}$. One merely applies Theorems 1–4 to $f=-F$ and $g=-G$.

REMARK 2: Let $A$, $B$ be positive definite and set $C=A+B$. Let $g(x_1,\ldots,x_m)=x_1^{-1}+\ldots+x_m^{-1}$. Writing $C=\frac{1}{2}(2A)+\frac{1}{2}(2B)$, we obtain from theorem 3 the inequality

$$4\sum_{s=1}^{m}\gamma_{k_s'}{}^{-1}\leq\sum_{s=1}^{m}\alpha_{i_s'}{}^{-1}+\sum_{s=1}^{m}\beta_{j_s'}{}^{-1}. \tag{10}$$

In particular if we take $i_s'=s$ and $j_s'=n-m+s$, for $s=1,\ldots,m$, then (10) becomes

$$4\sum_{s=1}^{m}\gamma_s^{-1}\leq\sum_{s=1}^{m}\alpha_s^{-1}+\sum_{s=1}^{m}\beta_{n-m+s}^{-1}. \tag{11}$$

If in theorem 4, we set $Z_1'=\ldots=Z_{n-m}'=m$, $W_1'=\ldots=W_{n-m}'=0$ and use this same $g$ we obtain (11). The inequality (11) is clearly sharper than the inequality

$$2\sum_{s=1}^{m}\gamma_s^{-1}\leq\sum_{s=1}^{m}\alpha_s^{-1}+\sum_{s=1}^{m}\beta_{n-m+s}^{-1}. \tag{12}$$

The inequality (12) was proved in [2] under the additional hypothesis $m\leq n/2$. Thus theorems 3 and 4 both substantially generalize and sharpen theorem 2 of [2].

By further specializations of theorems 1 and 2 many additional nonlinear inequalities may be written down. In particular many of the results in [3, chap. 2], [4], [10, p. 110] may be generalized.

REMARK 3: If in theorems 1 and 2, we let $A$, $B$, $C=\theta A+\phi B$ be not necessarily Hermitian matrices and we let the numbers (1) denote the singular values of $A, B, C$ respectively, then the conclusions of theorems 1 and 2 remain valid. This is because the inequalities (3) and (5) are known to be valid for the singular values of matrix sums. By this type of device many of the singular value inequalities in [1, chapter 2] and [4] may be be substantially generalized.

REMARK 4: The results of theorems 1 and 2 are used in [6].

## 3. A new inequality for the singular values of a matrix product.

THEOREM 5: *Let* A, B, C = AB *be a product of not necessarily Hermitian matrices. Let (1) deonote the singular values of* A, B, C *respectively. Let* $I_s$, $J_s$, $K_s$ *denote the integers defined in theorem 2. Then*

$$\sum_{s=1}^{m}\gamma_{K_s}\leq\sum_{s=1}^{m}\alpha_{I_s}\beta_{J_s}. \tag{13}$$

PROOF: The inequality

$$\prod_{s=1}^{m}\gamma_{K_s}\leq\prod_{s=1}^{m}\alpha_{I_s}\beta_{J_s}$$

is known [9]. As in the proof of theorem 2 it implies

$$\prod_{s=1}^{r}\gamma_{K_s}\leq\prod_{s=1}^{r}\alpha_{I_s}\beta_{J_s}$$

for each $r=1, \ldots, m$. *If* we set

$$c_s = log\, \gamma_{k_s}, \qquad d_s = log\, \alpha_{i_s}\beta_{j_s}$$

then $c_1+\ldots+c_r \leq d_1+\ldots+d_r$ for $r=1, \ldots, m$. The function $f(x_1, \ldots, x_m) = exp\, x_1+\ldots+\exp x_m$ is convex and increasing and hence $\exp c_1+\ldots+\exp c_m \leq exp\, d_1+\ldots+exp\, d_m$. This immediately yields (13). The inequality

$$\sum_{s=1}^{m} \gamma_{k_s} \leq \sum_{s=1}^{m} \alpha_{i_s}\beta_{j_s}$$

is also true and was proved by the same method in [9].

## 3. References


[1] Gohberg, I. C., and Krein, M. G., Introduction to the theory of linear nonselfadjoint operators, Transl. of Mathematical Monographs, Amer. Math. Soc., Vol. **18**, 1969, Chapter 2. MR39 #7447.

[2] Marcus, M., An inequality for linear transformations, Proc. Amer. Math. Soc., **18** (1967), 793–797. MR36 #3811.

[3] Marcus, M., and Minc, H., A Survey of Matrix Theory and Matrix Inequalities. Prindle, Weber, and Schmidt, 1969, Chapter 2. MR29 #112.

[4] Markus, A, S., The eigen- and singular values of the sum and product of linear operators, Russian Mathematical Surveys **19** (1964), 91–120. MR29 #6318.

[5] Ostrowski, A. S., Sur quelques applications des fonctions convexes et concaves au sens de I. Schur, J. Math. Pures Appl. **9** (31), 1952, 253–292. MR14, #625.

[6] Thompson, R. C., The real and absolute singular values of a matrix product, J. Linear Algebra and Appl., in press.

[7] Thompson, R. C., and Freede, L. J., Eigenvalues of sums of Hermitian matrices. J. Linear Algebra and Appl., in press.

[8] Thompson, R. C., Eigenvalues of sums of Hermitian matrices. II, Aequationes Math., **5**(1), 1970, 103–115.

[9] Thompson, R. C., and Therianos, S., The singular values of a matrix product. I, Scripta Math., in press.

[10] Zwahlen, B. P., Über die Eigenwerte der Summe zweier selbstadjungierter Operatoren, Comment. Math. Helv. **40** (1966), 81–116. MR32 #7566.


(Paper 75B3&4–351)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# A New Type of Boundary Value Coupling for Second Order Sturm-Liouville Systems

**Jon Geist**

**Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234**

**(October 6, 1971)**

A natural generalization of the familiar second order Sturm-Liouville system is presented. This generalization consists of considering a number of differential equations defined on different intervals, instead of just one equation on one interval. The self-adjoint character of the differential equations is retained in the generalization, but the boundary conditions are relaxed considerably. The most general boundary conditions which can be accommodated by this sort of generalization of Sturm-Liouville theory are discussed. The existence of eigenvalues is proved, and a generalized orthogonality and a weak eigenfunction expansion theorem are derived.

Key words: Boundary value problems; coupled equations; differential equations; self-adjoint equations; Sturm-Liouville systems.

## 1. Introduction

The solution to a large class of problems of interest in applied mathematics can be reduced to the solution of the second order Sturm-Liouville system,

$$(r(x)y'(x,\lambda))' = (q(x) - \lambda p(x))y(x,\lambda), \tag{1}$$

where $a \leq x \leq b$, and where

$$\alpha y(b,\lambda) + \beta y'(b,\lambda) = \gamma y(a,\lambda) + \delta y'(a,\lambda) = 0. \tag{2}$$

The primes indicate differentiation with respect to $x$, and the proper choice of $p(x) > 0$, $r(x) > 0$, $q(x)$, $a$, $b$, $\alpha$, $\beta$, $\gamma$, and $\delta$ is dictated by the particular problem. The theory of such Sturm-Liouville systems is well known [1, 2][1] and will not be reviewed here. Suffice it to say that it is a very powerful technique for handling the problems to which it is applicable.

Recently, generalizations of the second order Sturm-Liouville system have been developed which, while retaining the distinctive Sturm-Liouville character, are applicable to problems involving composite media [3, 4, 5, 6]. The purpose of this paper is to discuss a different generalization of the second order Sturm-Liouville system which retains the Sturm-Liouville character. In particular, an eigenvalue equation, an orthogonality theorem, and a weak eigenfunction expansion theorem will be derived. This eigenfunction expansion theorem is the first place where the results of this paper deviate from the previous work on Sturm-Liouville systems. It will be shown by counter-example, that the eigenfunctions of this Sturm-Liouville like system do not form a complete set in the sense of the previously described Sturm-Liouville systems.

The generalization of the second order Sturm-Liouville system discussed in this paper, and that which is applicable to composite media problems are similar in many ways. Each arises when

---

*AMS Subject Classification:* Primary 34B25; 42A64.

[1] Figures in brackets indicate the literature references at the end of this paper.

attempting to solve time dependent heat transfer problems in complicated systems, and each consists of considering a number of differential equations defined on different intervals, instead of just one equation defined on one interval. In both generalizations the self-adjoint character of the differential equations is retained, but the conditions which the solutions must satisfy at the ends of the intervals are relaxed. In both cases, the solutions to all of the equations are coupled by relations involving their values and the values of their first derivatives at the ends of the intervals on which they are defined. The difference between the two generalizations is in the specific form of the coupling. In the case of the composite media generalization, the coupling connects the boundary values at the right end of the first interval with those at the left end of the second interval; the boundary values at the right end of the second interval with those at the left end of the third interval, and so forth. In the alternate generalization, introduced here, the coupling connects the boundary values at the left ends of all of the intervals. Figures 1a and 1b show examples of the two types of coupling schematically. It is interesting that the form of coupling introduced here results in a simpler eigenvalue equation than that which results from the composite media generalization. In the case where only two intervals are considered, the composite media generalization is a special case of the generalization introduced here. Except for this case, this generalization of the second order Sturm-Liouville system does not appear to have been studied previously. An example of an application of the theory developed in this paper, is the calculation of the temperature distribution as a function of time in a number of wires whose left ends are all welded together and whose right ends are all connected to a heat sink.

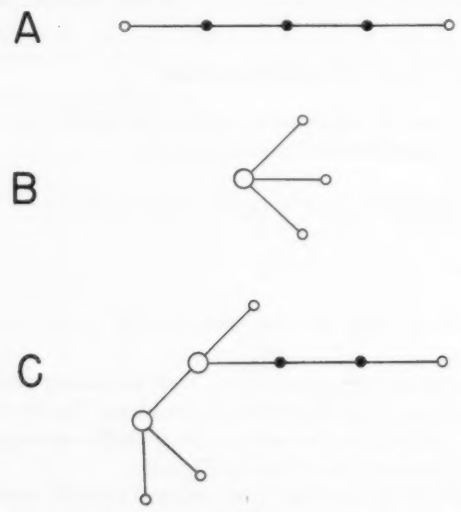


FIGURE 1. *Schematic representation of different types of Boundary value coupling*

– Interval on which self-adjoint differential equation is defined.
○ End point of interval where Sturm-Liouville boundary conditions are imposed.
● End point of interval where boundary conditions appropriate to the composite media generalization are imposed.
○ End point of interval where boundary conditions appropriate to the generalization introduced in this paper are imposed.

## 2. Postulates and Principal Results

Consider the $N$ self-adjoint differential equations obtained by letting $i=1, 2, \ldots, N$ in the equation,

$$(r_i(x_i)y_i'(x_i,\lambda))' = (q_i(x_i) - \lambda p_i(x_i))\, y_i(x_i,\lambda) \tag{3}$$

which is to hold for $a_i \leq x_i \leq b_i$, where $r_i(x_i) > 0$ and $p_i(x_i) > 0$. For the present, retain the usual Sturm-Liouville boundary condition,

$$\alpha_i y_i(b_i, \lambda) + \beta_i y_i'(b_i, \lambda) = 0, \tag{4}$$

at the right end of each interval, but at the left end of the intervals employ a generalized boundary condition, consisting of the $N-1$ equations obtained by letting $i = 1, 2, \ldots, N-1$ in the equation,

$$\gamma_i y_i(a_i, \lambda) + \delta_i y_i'(a_i, \lambda) = \gamma_{i+1} y_{i+1}(a_{i+1}, \lambda) + \delta_{i+1} y_{i+1}'(a_{i+1}, \lambda) \neq 0. \tag{5}$$

One more coupling condition is needed to give a total of 2N boundary conditions, so introduce the general relation,

$$\sum_{i=1}^{N} C_i y_i(a_i, \lambda) + D_i y_i'(a_i, \lambda) = 0, \tag{6}$$

where

$$D_i \gamma_i - C_i \delta_i > 0, \tag{7}$$

for $i = 1, 2, \ldots, N$.

Equations 3, 4, 5, 6 and 7 define the system. Now it is necessary to determine under what conditions a set of functions, $(y_1(x_1, \lambda), y_2(x_2, \lambda), \ldots, y_N(x_N, \lambda))$, exists such that these five equations are satisfied.

First concentrate attention on eqs (3) and (4). For any positive integer value of $i \leq N$, it can be shown, following Ince [7], that for every value of $\lambda$, eqs (3) and (4) determine, except for the specification of a multiplicative constant, one and only one solution which is continuous and has a continuous first derivative on the interval, $(a_i, b_i)$. Keeping this in mind, divide eq (6) by $\lambda_1 y_1(a_1, \lambda) + \delta_1 y'(a_1, \lambda)$ and substitute the $N-1$ equations represented by eq (5) into the result, to obtain the equation,

$$\sum_{i=1}^{N} \frac{C_i y_i(a_i, \lambda) + D_i y_i'(a_i, \lambda)}{\gamma_i y_i(a_i, \lambda) + \delta_i y_i'(a_i, \lambda)} = 0, \tag{8}$$

which is generally a transcendental equation in $\lambda$, and which is independent of the multiplicative constant which is as yet unspecified in $y_i(x_i, \lambda)$. Equation (8) is called the characteristic or eigenvalue equation of the system, since its roots are the eigenvalues of the problem; that is, the only values of $\lambda$ for which eqs (3), (4), (5), and (6) are mutually consistent. The existence of a countably infinite number of such roots is proved in appendix A of this paper. Let $\lambda_m$ be the $m$th root of eq (8), then it is called the $m$th eigenvalue of the system, and $y_i(x_i, \lambda_m)$ will be called the $m$th eigenfunction on the $i$th interval of the system. This completes the discussion of the eigenvalue equation.

Now attention will be turned to exhibiting the Sturm-Liouville like properties of the eigenfunctions. Other than the mere existence of eigenvalues, the most distinctive characteristic of Sturm-Liouville theory is the theorem expressing the orthogonality of the different eigenfunctions and the related eigenfunction expansion theorem. For the system described in this paper, the orthogonality theorem, which is proved in appendix B, takes the form,

$$\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i) y_i(x_i, \lambda_m) y_i(x_i, \lambda_n) dx_i = 0, \tag{9}$$

for $m \neq n$, where

$$W_i = \frac{D_i\gamma_i - C_i\delta_i}{r_i(a_i)} \tag{10}$$

Now the problem of an eigenfunction expansion theorem will be considered. Since the functions, $y_i(x_i, \lambda_m)$, are determined only to within a multiplicative constant, it is convenient to define

$$d_{im} = \left[\int_{a_i}^{b_i} p_i(x_i) y_i(x_i, \lambda_m) y_i(x_i, \lambda_m) dx_i\right]^{1/2}, \tag{11}$$

and

$$\phi_{im}(x_i) = y_i(x_i, \lambda_m)/d_{im}. \tag{12}$$

Thus each of the functions, $\phi_{im}(x_i)$, is square normalized to unity on the $i$th interval with respect to the weighting functions, $p_i(x_i)$, and this property uniquely determines the function, $\phi_{im}(x_i)$. Of course, the $d_{im}$ are as yet undetermined, since the $y_i(x_i, \lambda_m)$ are determined only to within a multiplicative constant.

Now consider the possibility of choosing these multiplicative constants (and hence the $d_{im}$) in such a way as to express the $N$ arbitrary functions, $g_i(x_i)$, $i = 1, 2, \ldots, N$, as a sum of eigenfunctions. For any given value of $i$, the constants, $d_{im}$, which are obtained by varying $m$ are independent of each other. However, when $m$ is held constant and $i$ is varied, the resulting $d_{im}$ are not independent. Once one, say $d_{jm}$, is specified all of the others can be obtained from the equation,

$$\frac{d_{im}}{d_{jm}} = \frac{\gamma_j\phi_{jm}(a_j) + \delta_j\phi'_{jm}(a_j)}{\gamma_i\phi_{im}(a_i) + \delta_i\phi'_{im}(a_i)}, \tag{13}$$

which is obtained by substituting eq (12) into eq (5). Thus in the eigenfunction expansions,

$$f_i(x_i) = \sum_{m=1}^{\infty} d_{im}\phi_{im}(x_i), \tag{14}$$

the $N$ functions, $f_i(x_i)$, cannot be chosen arbitrarily. Once one of them is specified, all of the rest of them are uniquely determined. This is a strong restriction, but it does not necessarily prevent the theory from being useful. However, it does weaken the eigenvalue expansion theorem considerably. The theorem stated below and proved in appendix C is the strongest that can be proved when the functions to be expanded are left completely arbitrary.

For each value of $j = 1, 2, \ldots, N$, the function,

$$f_j(x_j) = \sum_{m=1}^{\infty} d_{jm}\phi_{jm}(x_j) = \sum_{m=1}^{\infty} y_j(x_j, \lambda_m), \tag{15}$$

will be a "best" approximation to the function, $g_j(x_j)$, in the sense that the expression,

$$I = \sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)[g_i(x_i) - f_i(x_i)]^2 dx_i, \tag{16}$$

is a minimum, if and only if for each $m = 1, 2, \ldots,$

$$d_{jm} = \frac{\sum_{i=1}^{N} R_{mji} W_i \int_{a_i}^{b_i} p_i(x_i) g_i(x_i) \phi_{im}(x_i) dx_i}{\sum_{i=1}^{N} (R_{mji})^2 W_i} \tag{17}$$

where

$$R_{mji}=\frac{d_{im}}{d_{jm}}=\frac{\gamma_j\phi_{jm}(a_j)+\delta_j\phi'_{jm}(a_j)}{\gamma_i\phi_{im}(a_i)+\delta_i\phi'_{im}(a_i)}, \tag{18}$$

Since the above eigenfunction expansion theorem does not require that each function, $f_i(x_i)$, be identical to the function, $g_i(x_i)$, a criterion for goodness of fit is useful. It is shown in appendix C that the ratio,

$$R=\sum_{n=1}^{\infty}\sum_{i=1}^{N}W_i(d_{in})^2 \Big/ \sum_{i=1}^{N}W_i\int_{a_i}^{b_i}p_i(x_i)[g_i(x_i)]^2dx_i \tag{19}$$

has the property that, $0\leq R\leq 1$, and $R=1$ if and only if $f_i(x_i)=g_i(x_i)$, for $i=1, 2, \ldots, N$. Thus $R$ can be used as a criterion for goodness of fit in a qualitative manner. Furthermore the set of functions for which the eigenfunctions of the problem form a complete set are characterized by $R=1$.

### 3. An Example

As an example, let $i=1$, 2, and 3, in eqs (3), (4), (5), and (6), and let $a_i=-1$, $b_i=0$, $\epsilon=\lambda^{1/2}$, $r_i(x_i)=1$, $q_i(x_i)=0$, $p_i(x_i)=(i)^2$, $\alpha_i=1$, $\beta_i=0$, $\gamma_i=1$, $\delta_i=0$, $C_i=0$, and $D_i=1$. The resulting system arises when attempting to determine the temperature distribution, $T_i(x_i, t)$, as a function of time in three straight wires of identical thermal conductivity and area, whose specific heats per unit volume are in the ratio 1, 4, 9, and whose right ends are attached to a block whose thermal conductivity and heat capacity are infinite, and whose left ends are attached to a block whose thermal conductivity is infinite, but whose heat capacity is zero. If the initial condition for this problem is that $T_i(x_i, 0)=-x_i$, for $i=1, 2, 3$, then $T_i(x_i, t)$ can be written in the form

$$T_i(x_i, t)=\sum_{m=1}^{\infty}d_{im}\phi_{im}(x_i)\exp\left[-(\epsilon_m)^2t\right], \tag{20}$$

where $\epsilon_m^2=\lambda_m$ is the $m$th root of eq (8), if and only if $f_i(x_i)=-x_i$, for $i=1, 2, 3$, where $f_i(x_i)$ is defined in eq (15). However, this is not the case. The functions,

$$Z_i(x_i)=\sum_{m=1}^{150}y_i(x_i, \lambda_m), \tag{21}$$

are shown in figure 2. It is clear from the figure, that these functions are not very good approximations to $-x_i$. For this particular case, $R=0.990159$. There is no significant difference between these results and the results when only the first fifteen terms are included in the sum, and presumably there is no significant difference between these results and the results when all of the terms are included in the sum. Thus in order to obtain $T_i(x_i, t)$ it is necessary to add to the right hand side of eq (20), a nonseparable function of $x_i$ and $t$, which is a solution to the corresponding time dependent heat transfer equation for the problem. The properties of such non-separable solutions remain to be investigated.

Now suppose that we eliminate the third interval in the above problem, restricting $i$ to be 1 or 2. Then the problem is equivalent to a composite media problem, and interestingly enough the eigenfunctions form a complete set as can be seen from figure 3. In this case, $R=0.999989$, for just the first fifteen terms of the sum.

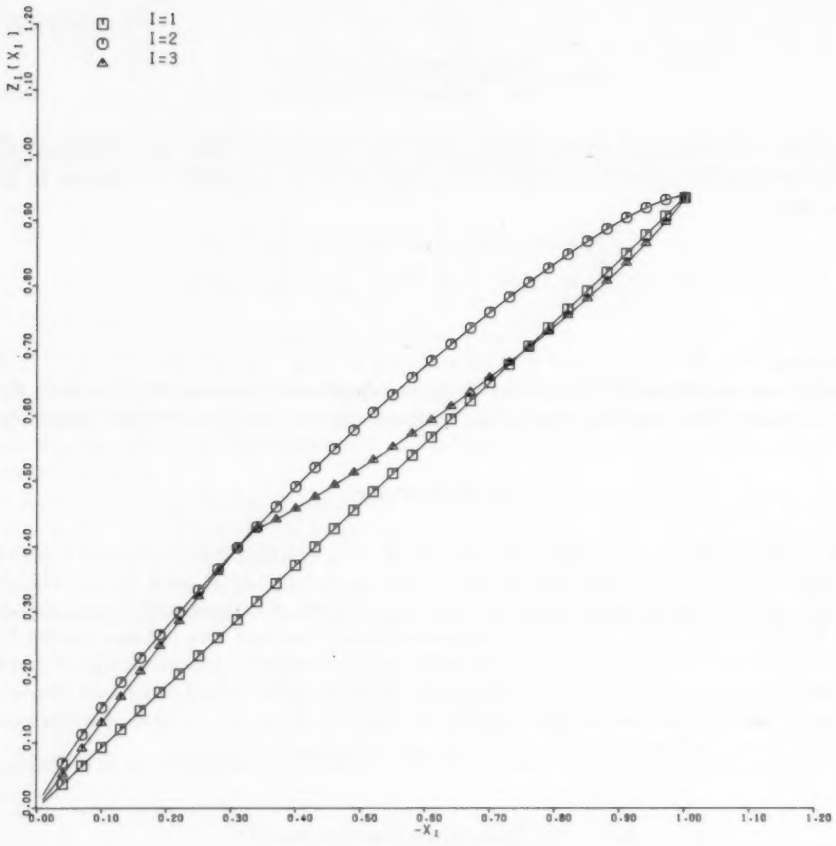


FIGURE 2. *Graphs of $Z_i(x_i)$ for $i=1$, 2, and 3, showing that $Z_i(x_i)$ is not a very good approximation to $-x_i$ in the three interval example considered in the paper.*

## 4. Discussion

The fact that nonseparable solutions appear in a prominent way in the generalization of Sturm-Liouville theory discussed here certainly limits the usefulness of the theory to some extent. How serious this limitation is in various applications will depend upon the general properties of these non-separable solutions. Thus an investigation of such non-separable functions would appear to be quite worthwhile.

Nowhere in the derivation of the results presented in this paper does the assumption that $a_i < x_i < b_i$ play a fundamental role. All of the results would have been identical if for some of the values of $i$, $b_i < x_i < a_i$, but the discussion would have been somewhat more complicated. To further generalize the second order Sturm-Liouville system, the boundary conditions at the points, $b_i$, could be relaxed. It is possible to replace them by generalized boundary conditions identical in form to those stated earlier in this paper, by introducing a number of new intervals on which more self-adjoint differential equations are defined. The type of coupling envisioned here should be obvious after glancing at figure 1c. The detailed development of this further generalization of the second order Sturm-Liouville system introduces no new ideas, and it is fraught with notational difficulties. Consequently it will not be pursued here. However, a few remarks are in order.

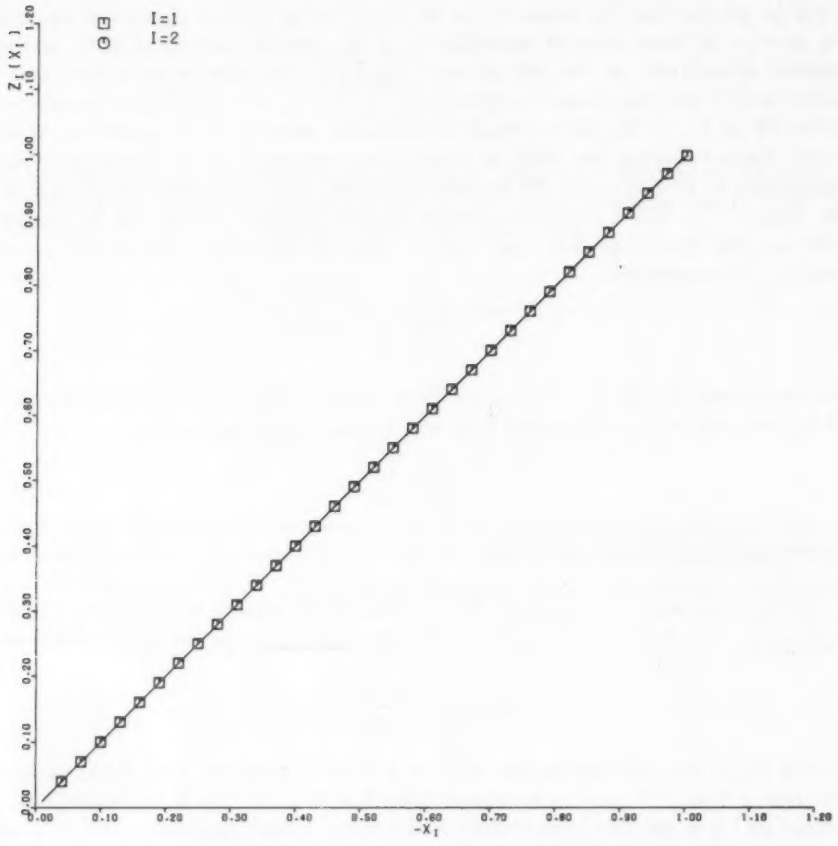


FIGURE 3. *Graphs of* $Z_i(x_i)$ *for* $i=1$, *and 2, showing that* $Z_i(x_i)$ *is an excellent approximation to* $-x_i$ *in the two interval (composite media) example considered in the paper.*

First, cross-linking among the branching intervals will not generally allow a nontrivial solution. But it is possible to cross-link any two intervals for which $r(a)=r(b)$ for both intervals. Second, both the composite media generalization and the generalization described earlier in this paper are special cases of this generalization of the second order Sturm-Liouville system. Third, in a detailed development of this theory, it will not be possible to derive a simple eigenvalue equation as was done here. A determinantal equation, similar in principle to, but more complicated in actual detail than that in reference [4] will have to be derived.

## 5. Appendix A

In this appendix, eq (8) will be proved to have an infinite number of roots. In discussing eq (8), it is convenient to define,

$$F_i(\lambda)=\frac{C_i\gamma_i(a_i,\lambda)+D_i\gamma_i'(a_i,\lambda)}{\gamma_i\gamma_i(a_i,\lambda)+\delta_i\gamma_i'(a_i,\lambda)}, \tag{1A}$$

and,

$$F(\lambda)=\sum_{i=1}^{N}F_i(\lambda). \tag{2A}$$

First it will be proved that for each $i = 1$ to $N$, $F_i(\lambda)$ has an infinite number of singularities, and that it is positive in some interval immediately to the right of each singularity, and negative in some interval immediately to the left of each singularity. In order to prove this, it is necessary to consider each of the four cases, $C_i = \delta_i = 0$, $D_i = \gamma_i = 0$, $C_i\gamma_i \neq 0$, and $D_i\delta_i \neq 0$, separately. Every possible choice of $C_i$, $\gamma_i$, $D_i$, and $\delta_i$ which is consistent with eq (7) is covered by at least one of these cases. Before looking into each of these cases separately, it is necessary to review some of the properties of $y_i(a_i, \lambda)$. Ince [8] has shown that if $y_i(x_i, \lambda)$ satisfies eqs (3) and (4), then the equation, $y_i(a_i, \lambda) = 0$, has a countably infinite number of roots. These will be denoted by $\mu_{im}$ where the $\mu_{im}$ are labeled so that $\mu_{im} < \mu_{im+1}$. Ince [9] has shown that as $\lambda$ is increased from $\mu_{im}$ to $\mu_{im+1}$, the expression,

$$\frac{y_i'(a_i, \lambda)}{y_i(a_i, \lambda)},$$

decreases monotonically from $+\infty$ to $-\infty$. Thus $y_i'(a_i, \lambda)$ also has an infinite number of roots, denoted by $\omega_{im}$, and as $\lambda$ is increased from $\omega_{im}$ to $\omega_{im+1}$, the expression,

$$\frac{y_i(a_i, \lambda)}{y_i'(a_i, \lambda)},$$

increases monotonically from $-\infty$ to $+\infty$.

Case 1: $C_i = \delta_i = 0$

In this case,

$$F_i(\lambda) = \frac{D_i y_i'(a_i, \lambda)}{\gamma_i y_i(a_i, \lambda)}, \tag{3A}$$

and dividing eq (7) by $(\gamma_i)^2$ shows that $D_i/\gamma_i$ is a positive number. Thus $F_i(\lambda)$ decreases monotonically from $+\infty$ to $-\infty$, as $\lambda$ is increased from $\mu_{im}$ to $\mu_{im+1}$, so it has an infinite number of singularities and it is positive immediately to the right of each singularity and it is negative immediately to the left of each singularity.

Case 2: $D_i = \gamma_i = 0$

In this case,

$$F_i(\lambda) = \frac{C_i}{\delta_i} \frac{y_i(a_i, \lambda)}{y_i'(a_i, \lambda)}, \tag{4A}$$

but dividing eq (7) by $(\delta_i)^2$ shows that $C_i/\delta_i$ is a negative number. Thus $F_i(\lambda)$ decreases monotonically from $+\infty$ to $-\infty$ as $\lambda$ is increased from $\omega_{im}$ to $\omega_{im+1}$. So again, $F_i(\lambda)$ has an infinite number of singularities, is positive to the right of each singularity, and is negative to the left of each singularity.

Case 3: $D_i\delta_i \neq 0$

$$\text{Let } f_i(\lambda) = \left(\frac{C_i}{D_i} + \frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}\right)\left(\frac{\gamma_i}{\delta_i} + \frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}\right)^{-1}, \tag{5A}$$

then $F_i(\lambda) = (D_i/\delta_i) f_i(\lambda)$. Now the equations,

$$\frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)} + \frac{C_i}{D_i} = 0, \tag{6A}$$

and

$$\frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}+\frac{\gamma_i}{\delta_i}=0, \tag{7A}$$

each have one and only one root between $\mu_{im}$ and $\mu_{im+1}$. This root of eq (7A) will be denoted by $\xi_{im}$ and the equivalent root of eq (6A) will be denoted by $\eta_{im}$. First supposing that $0<D_i\delta_i$, use eq (7) to show that $-\gamma_i/\delta_i<-C_i/D_i$, so that $\mu_{im}<\eta_{im}<\xi_{im}<\mu_{im+1}$. Thus

$$\frac{-\gamma_i}{\delta_i}<\frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}<-\frac{C_i}{D_i} \text{ for } \eta_{im}<\lambda<\xi_{im}, \tag{8A}$$

and

$$\frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}<-\frac{\gamma_i}{\delta_i} \text{ for } \xi_{im}<\lambda<\mu_{im+1}. \tag{9A}$$

Since $\xi_{im}$ is a root of eq (7A), $f_i(\lambda)$ has a singularity at $\xi_{im}$, and eq (8A) can be used to show that $f_i(\lambda)$ is negative when $\eta_{im}<\lambda<\xi_{im}$, and eq (9A) can be used to show that $f_i(\lambda)$ is positive when $\xi_{im}<\lambda<\mu_{im+1}$. Letting $m=1,2,\ldots$ in the above discussion shows that $f_i(\lambda)$ has an infinite number of singularities and that it is positive immediately to the right of each singularity and that it is negative immediately to the left of each singularity. Now, supposing that $D_i\delta_i<0$, use eq (7) to show that

$$\frac{-C_i}{D_i}<\frac{-\gamma_i}{\delta_i},$$

so that $\mu_{im}<\xi_{im}<\eta_{im}<\mu_{im+1}$.

Thus

$$-\frac{C_i}{D_i}<\frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}<-\frac{\gamma_i}{\delta_i} \text{ for } \xi_{im}<\lambda<\eta_{im}, \tag{10A}$$

and

$$-\frac{\gamma_i}{\delta_i}<\frac{y_i'(a_i,\lambda)}{y_i(a_i,\lambda)}, \text{ for } \mu_{im}<\lambda<\xi_{im} \tag{11A}$$

Since $\xi_{im}$ is a root of eq (7A), $f_i(\lambda)$ has a singularity at $\xi_{im}$, and eq (10A) implies that $f_i(\lambda)$ is negative when $\xi_{im}<\lambda<\eta_{im}$, and eq (11A) implies that $f_i(\lambda)$ is positive when $\mu_{im}<\lambda<\zeta_{im}$. Letting $m=1, 2, \ldots$ in the above discussion shows that $f_i(\lambda)$ has an infinite number of singularities and it is negative in some interval immediately to the right of each singularity and that is positive in some interval immediately to the left of each singularity. So in both the case where $D_i\delta_i>0$ and in the case where $D_i\delta_i<0$, $F_i(\lambda)=(D_i/\delta_i)\,f_i(\lambda)$ has an infinite number of singularities and it is positive immediately to the right of each singularity and negative immediately to the left of each singularity.

Case 4: $C_i\gamma_i\neq 0$

The proof is similar at each step to that of Case 3 and will not be presented here.

If $F_i(\lambda)$ has a singularity at $\kappa_{im}$, then eq (2A) implies that $F(\lambda)$ also has a singularity there, and $F(\lambda)$ behaves like $F_i(\lambda)$ in some neighborhood of $\kappa_{im}$. Thus, if $\kappa_{im}$ and $\kappa_{jn}$ are two consecutive



449-764 OL - 72 - 3

singularities of $F(\lambda)$, and $\kappa_{im} < \kappa_{jn}$, then $F(\lambda)$ is positive immediately to the right of $\kappa_{im}$ and negative immediately to the left of $\kappa_{jn}$. Now since $F(\lambda)$ is continuous for $\kappa_{im} < \lambda < \kappa_{jn}$, it must have a zero in this interval. Thus eq (8) has at least one root between any two consecutive singularities of $F(\lambda)$. Finally, since $F(\lambda)$ has an infinite number of singularities, it has an infinite number of roots, as was to be proved.

## 6. Appendix B

In this appendix, eq (9) will be proved. The proof requires a lemma which will be proved separately, but in somewhat more general form than needed.

### Lemma

Since this lemma consists almost entirely of algebraic manipulations, it will prove convenient to use the shorthand notation,

$$M_i = M_i(a_i, \xi), \tag{1B}$$

$$N_i = N_i(a_i, \eta), \tag{2B}$$

and

$$r_i = r_i(a_i), \tag{3B}$$

where $M_i(x_i, \xi)$ and $N_i(x_i, \eta)$ satisfy eqs (3), (4), and (5), with $\lambda = \xi$, and $\lambda = \eta$, respectively. Now let

$$S = (\gamma_i N_i + \delta_i N_i') \sum_{j=1}^{N} (C_j M_j + D_j M_j') - (\gamma_i M_i + \delta_i M_i') \sum_{j=1}^{N} (C_j N_j + D_j N_j') , \tag{4B}$$

and notice that if $M_j$ and $N_j$ both satisfy eq (6), then $S = 0$, but if only one of them satisfies eq (6), then $S \neq 0$.

Since both $M_i$ and $N_i$ satisfy eq (5),

$$S = \sum_{j=1}^{N} [(C_j M_j + D_j M_j')(\gamma_j N_j + \delta_j N_j') - (C_j N_j + D_j N_j')(\gamma_j M_j + \delta_j M_j')] \cdot \tag{5B}$$

When the indicated multiplications in eq (5B) are carried out, some of the terms cancel one another, and the remaining terms can be combined to yield,

$$S = \sum_{j=1}^{N} (D_j \gamma_j - C_j \delta_j)(M_j' N_j - N_j' M_j) \cdot \tag{6B}$$

Thus by eq (10),

$$S = \sum_{j=1}^{N} W_j r_j (M_j' N_j - N_j' M_j) \tag{7B}$$

Churchill [10] has shown that eqs (3) and (4) are sufficient conditions for

$$(\xi - \eta) \int_{a_j}^{b_j} p_j(x_j) M_j(x_j, \xi) N_j(x_j, \eta) dx_j = r_j (M_j' N_j - N_j' M_j). \tag{8B}$$

Thus

$$(\xi-\eta)\sum_{j=1}^{N} W_j \int_{a_j}^{b_j} p_j(x_j)M_j(x_j,\xi)N_j(x_j,\eta)dx_j=S, \tag{9B}$$

where $S$ is given by eq (4B). This is the significant result of the lemma.

Now let $\xi=\lambda_m$ and $\eta=\lambda_n$. Then $M_j(x_j, \xi)=y_j(x_j, \lambda_m)$ and $N_j(x_j, \eta)=y_j(x_j, \lambda_n)$, so $M_j$ and $N_j$ both satisfy eq (6), so $S=0$. And if $m \neq n$, eq (9B) becomes eq (9). This completes the proof.

## 7. Appendix C

In this appendix, it will be shown that eq (17) is a necessary and sufficient condition for the expression, $I$, defined in eq (16), to have a minimum with respect to variation of the $d_{jm}$ subject to the constrants imposed by the $N-1$ equations represented by eq (13).

First notice that by the restrictions stated after eq (3), by eq (7) and eq (10), $I \geq 0$, so it has a minimum. Also notice that $I=0$, if and only if $f_i(x_i)=g_i(x_i)$ for $i=1, 2, \ldots, N$. Next take the variation of $I$ with respect to $d_{jm}$ where $m=1, 2, \ldots$, and $j$ is fixed. Using eq (13) to justify replacing the partial derivative of $d_{im}$ with respect to $d_{jm}$ by $d_{im}/d_{jm}$, rearrange the result to obtain,

$$\delta(I)=\sum_{m=1}^{\infty} \delta(d_{jm}) \sum_{i=1}^{N} \left(\frac{d_{im}}{d_{jm}}\right) W_i \int_{a_i}^{b_i} p_i(x_i)\,(g_i(x_i)-f_i(x_i))\phi_{im}(x_i)dx_i=0. \tag{1C}$$

Now since each $d_{jm}$ is independent of $d_{jn}$ for $n \neq m$, the coefficient of every $\delta(d_{jm})$ must be zero. Combining eqs (15), and (18) with this result yields

$$\sum_{i=1}^{N} R_{mji}W_i \int_{a_i}^{b_i} p_i(x_i)g_i(x_i)\phi_{im}(x_i)dx_i=\frac{1}{d_{jm}}\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,f_i(x_i)y_i(x_i, \lambda_m)dx_i, \tag{2C}$$

for $m=1, 2, \ldots$. Finally applying eqs (15), (9), and (11) to the right-hand side of eq (2C), and again using eq (18) one sees that the simultaneous equations are diagonalized and that they are equivalent to eq (17). All of the steps in this proof are reversible, so eq (17) is a necessary and sufficient condition for $I$ to have a minimum value as was to be shown.

In order to obtain a criterion for goodness of fit, expand $I$ to obtain,

$$I=\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,(g_i(x_i))^2dx_i+\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,(f_i(x_i))^2dx_i-2\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)g_i(x_i)f_i(x_i)dx_i. \tag{3C}$$

Notice that if eq (2C) is multiplied by $d_{jm}$ and summed over $m$, the result is

$$\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)g_i(x_i)f_i(x_i)dx_i=\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,(f_i(x_i))^2dx_i. \tag{4C}$$

Substitute this result into eq (3C) to obtain

$$I=\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,(g_i(x_i))^2dx_i-\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,(f_i(x_i))^2dx_i. \tag{5C}$$

Finally substitute eq (15) into eq (5C), and use the orthogonality theorem, eq (9) to obtain,

$$I=\sum_{i=1}^{N} W_i \int_{a_i}^{b_i} p_i(x_i)\,(g_i(x_i))^2dx_i-\sum_{m=1}^{\infty}\sum_{i=1}^{N} W_i(d_{im})^2. \tag{6C}$$

Since $I \geq 0$, $0 \leq R \leq 1$, where $R$ was defined in eq (19). The second property of $R$ follows from the property of $I$ that it is zero if and only if $g_i(x_i) = f_i(x_i)$ for $a_i \leq x_i \leq b_i$, and $i = 1, 2, \ldots, N$.

---

I thank R. C. Kraft and W. P. Reid for constructive criticism on various aspects of this work, and J. S. Gallagher for preparing some of the figures.

## 8. References

[1] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., p. 217 and Ch. 10.

[2] Churchill, R. V., Operational Mathematics (McGraw-Hill, New York, 1958, 2nd ed., Ch. 9.

[3] Reid, W. P., J. Franklin Inst. **274,** 352 (1962).

[4] Bulavin, P. E. and Kashcheev, V. M., Int. Chem. Eng., **5,** 112 (1965).

[5] Tittle, C. W., J. Appl. Phys. **36,** 1486 (1965).

[6] Formerly such problems were solved by using integral transforms, and the evaluation of a complex inversion integral was required. See Carslaw, H. C. and Jaeger, J. C., Conduction of Heat in Solids, (Oxford University Press, London, 1959), 2d ed., p. 319, and references listed there. Actually, the two techniques are not unrelated. An orthogonality theorem identical to that employed in the direct eigenvalue expansion is often used to study the properties of solutions derived by transform techniques, and the same eigenvalue equation occurs in both cases.

[7] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., pp. 73–75.

[8] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., pp. 231–232.

[9] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., pp. 229–230, 232.

[10] Churchill, R. V., Operational Mathematics (McGraw-Hill, Inc., New York, 1958), 2d ed., pp. 264–265. (Note that Churchill denotes by $-\lambda$, what Ince and the author of the present paper denote by $\lambda$.)

(Paper 75B3&4–352)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# An Ergodic Source Which Is not a Function of Any Finite Markov Process*

**Roger D. Traub****

**Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234**

**(June 19, 1971)**

The problem of deciding when a stationary stochastic process is a function of a finite Markov process has been considered by several authors, leading to an elegant necessary and sufficient condition. Taking a different approach, this note uses elementary ring theory to prove that a certain explicitly constructed stationary ergodic process is not a function of any finite Markov process.

Key words: Information theory; Markov process; probability theory; ring theory; stochastic process.

## 1. Introduction

The problem of deciding when a stationary stochastic process is a function of a finite Markov process has been considered by several authors ([1, 2, 4–6]),[1] leading to an elegant necessary and sufficient condition (Heller, [6]). This condition involves both (a) the finite dimensionality of a certain module over the real numbers, and (b) the requirement that a certain cone in this module be polyhedral (an idea used earlier by Dharmadhikari [2]). The output symbols of the stochastic process act as linear operators on the above-mentioned module.

Taking a different approach, this note uses elementary ring theory to prove that a certain specific stationary ergodic process is not a function of any finite Markov process. The structure of this process—although not the idea—resembles that of an example of Fox and Rubin [4].

Section 2 presents the principal tool for the verification, namely a sufficient condition for a "source" (in the language of information theory) *not* to be a function of any finite Markov process. This condition and its verification are purely algebraic.

Section 3 gives the explicit description of the example, plus the verification of its properties. A related conjecture is posed in section 4.

The balance of the present section 1 consists of definitions and other background material provided to make the exposition more nearly self-contained.

Let A be a fixed finite nonempty set; the sources to be considered emit "symbols" from this "alphabet" $A$ as outputs. Let $A^\infty$ be the set of all doubly infinite sequences $\mathbf{x}=\{x_t\}_{-\infty}^{\infty}$ of members of $A$; each such sequence can be considered a possible time history of the source's outputs, with symbol $x_t$ emitted at time $t$. The *shift* operator $T: A^\infty \to A^\infty$ is defined by

$$T\mathbf{x}=\mathbf{y} \text{ where } y_t=x_{t+1}. \tag{1.1}$$

A *cylinder set* is a subset of $A^\infty$ of the form

$$C(J, g)=\{\mathbf{x}\epsilon A^\infty: x_t=g(t) \text{ for all } t\epsilon J\}$$

---

*AMS Subject Classification:* Primary 6285.

* Helpful discussions with A. J. Goldman, J. Levy and M. Newman are gratefully acknowledged.

** National Bureau of Standards and University of Pennsylvania Medical School.

[1] Figures in brackets indicate the literature references at the end of this paper.

where $J$ is some finite subset of the integers (positive, negative, or zero) and $g: J \to A$ is some function. If $J = \{1, 2, \ldots, k\}$ for some $k$, then $C(J, g)$ is called an *elementary* cylinder set. It follows from the finiteness of $A$ that every cylinder set is a finite union of elementary cylinder sets.

Let $\Sigma$ be the Borel field generated by all cylinder sets (or equivalently, all elementary cylinder sets). We now *define* an information source over $A$ to be a probability measure $\mu$ on $(A^\infty, \Sigma)$. Such a source will be denoted $[A, \mu]$. In terms of the notation in (1.1), the source is called *stationary* if $\mu(TS) = \mu(S)$ for every measurable subset $S$ of $A^\infty$, and is called *ergodic* if $TS = S$ (where $S$ is measureable) implies $\mu(S) = 0$ or $\mu(S) = 1$.

Let $I_n = \{1, 2, \ldots, n\}$, and suppose there is a function $f: I_n \to A$, and an $n$-state Markov chain with transition-matrix entries $p(i, j)$ and stationary probabilities $p(i)$, such that, for *every* elementary cylinder set

$$C = C(\{1, 2, \ldots, k\}, g),$$

we have

$$\mu(C) = \Sigma p(i_1) p(i_1, i_2) \ \ldots \ p(i_{k-1}, i_k) \tag{1.2}$$

where the sum is over all sequences $\{i_1, i_2, \ldots, i_k\}$ over $I_n$ such that $f(i_t) = g(t)$ for $t = 1, 2, \ldots, k$. Note that this entirely determines $\mu$, so that $[A, \mu]$ is specified by the finite set of data $\{p(i), p(i, j)\}$ together with $f$. In this case, $[A, \mu]$ is called a *function of a* (finite) *Markov process;* it is automatically stationary, and is ergodic if the Markov chain is.

If $I_n$, in the last paragraph, is replaced by the set of all natural numbers, we are led instead to the concept of "function of a denumerable-state Markov process." The example to be given later falls into this category, but not into the previous one, as will be proven.

## 2. An Algebraic Condition

Let $\Pi_N$ denote the set of all polynomials in $N$ variables with integer coefficients. For the information source $[A, \mu]$, define the following subring of the ring of real numbers:

$$V(A, \mu) = \{\rho : \rho = P(\mu(C_1), \ldots, \mu(C_N)) \text{ for some } N > 0, \text{ some } P \epsilon \Pi_N,$$

$$\text{and some cylinder sets } C_1, C_2, \ldots, C_N\}.$$

Suppose in particular that $[A, \mu]$ is a function of an $n$-state Markov chain with parameters $\{p(i), p(i, j)\}$. Then it follows by (1.2) that the measure of any elementary cylinder set is the value, at $(p(1), \ldots, p(n), p(1, 1), \ldots, p(n, n))$ of a polynomial in $n + n^2$ variables with integer coefficients. It follows that the same is true of the measure of *any* cylinder set, and so is true of every member of $V(A, \mu)$. That is, if we set

$$W_N(u_1, \ldots, u_N) = \{\rho : \rho = P(u_1, \ldots, u_N) \quad \text{for some} \quad P \epsilon \Pi_N\},$$

then we have,

$$V(A, \mu) \subseteq W_{n+n^2}(p(1), \ldots, p(n), p(1, 1), \ldots, p(n, n))$$

whenever $[A, \mu]$ is a function of the $n$-state Markov chain with the indicated parameters.

This observation provides the basis for the following result.

THEOREM 1: *Suppose there are real numbers* $\{\rho_i\}_1^m$ *such that the ring generated over the integers by* $V(A, \mu) \cup \{\rho_i\}_1^m$ *includes all the rationals. Then the source* $[A, \mu]$ *is not a function of any finite Markov process.*

This will be shown to follow from the next theorem, whose conclusion, in order to support an inductive proof, is stronger than is needed for Theorem 1.

THEOREM 2: *Let* $U=(u_1, \ldots, u_N)$ *be a fixed* N*-tuple of real numbers,* k *a fixed positive integer,* $(D_1, \ldots, D_k)$ *fixed polynomials in* $\Pi_N$ *with each* $D_j(U) \neq 0$. *Then the set* $R_N=\{\rho : \rho = P(U)/\prod_1^k [D_j(U)]^{a(j)}$ *for some* $P \epsilon \Pi_N$ *and some integral* $a(j) \geq 0\}$ *does not include all the rationals.*

We first show that Theorem 2 implies Theorem 1. In Theorem 2 let $N=m+n+n^2$, let $U=(\rho_1, \ldots, \rho_m, p(1), \ldots, p(n), p(1, 1), \ldots, p(n, n))$, and let all $D_j \equiv 1$. It follows from (2.1) that, if $[A, \mu]$ is a function of the Markov process with the indicated parameters $\{p(i), p(i, j)\}$, then $V(A, \mu) \subseteq R_N$. Since $\rho_i \epsilon R_N$, and $R_N$ is a ring over the integers, the ring $R^*$ generated over the integers by $V(A, \mu) \cup \{\rho_i\}_1^m$ also lies in $R_N$. By Theorem 2, $R^*$ cannot contain all rationals, and so Theorem 1 is proved. It is interesting that the proof uses only (1.2), and not the further properties

$$p(i, j) \geq 0, \qquad p(i) \geq 0,$$

$$\sum_j p(i, j) = 1, \qquad \sum_i p(i) p(i, j) = p(j).$$

We turn now to the proof of Theorem 2, based on a suggestion by M. Newman. The notation $R[z_1, \ldots, z_n]$ will mean the ring of polynomials in the indeterminates $z_1, \ldots, z_n$ with coefficients in a ring $R$. The proof is by induction on $N$.

CASE 1: $N=1$. Suppose the conclusion is false, i.e., that there is $U=u_1$ such that $R_1$ includes all the rationals. Let $I$ denote the ring of integers.

$U=u_1$ must be algebraic over the rationals, since for any prime $p$, $1/p \epsilon R_1$, which implies an equation of the form

$$\prod_1^k [D_j(U)]^{a(j)} - pP(U) = 0$$

in which not all coefficients can vanish if $p$ is chosen so as not to divide all coefficients of the polynomial $\prod_1^k D_j$. Now let $u_i = u_1^{(1)}$ and let $u_1^{(i)}$ $(2 \leq i \leq m)$ be the $i$th conjugate of $u_1$ over the rationals, so that $\prod_1^m (z - u_1^{(i)})$ is the minimal polynomial of $u_1$ over the rationals. Then for a suitable integer $\lambda$,

$$g(z) = \lambda \prod_1^m (z - u_1^{(i)}) \epsilon \Pi_1.$$

Let $\{\sigma_i\}_1^m$ denote the elementary symmetric functions of $m$ variables, and put

$$s_i = \sigma_i(u_1^{(1)}, \ldots, u_1^{(m)}),$$

so that

$$g(z) = \lambda (z^m - s_1 z^{m-1} + \ldots + (-1)^m s_m).$$

It follows that all $\lambda s_i \epsilon I$.

Let $p$ be a prime, to be determined later. As above, $1/p \epsilon R_1$ leads to an equation $f(u_1)=0$ where

$$f \equiv \prod_1^k (D_j)^{a(j)} - pp$$

for some appropriate $P \epsilon \Pi_1$ and nonnegative integers $a(j)$. Since $g$ is a constant multiple of the minimal polynomial of $u_1$, $f(u_1)=0$ implies that $f=gh$ for some polynomial $h$ with rational coefficients. Since $g(u_1^{(i)})=0$ for all $i$, all $f(u_1^{(i)})=0$ and so

$$1/p = P(u_1^{(p)})/\Pi_1^k [D_j(u_1^{(p)})]^{a(j)},$$

implying

$$1/p^m = \Pi_1^m P(u_1^{(p)})/\Pi_{i=1}^m \Pi_{j=1}^k [D_j(u_1^{(p)})]^{a(j)}. \tag{2.2}$$

In the above note that

$$\Pi_1^k [D_j(u_1^{(p)})]^{a(j)}$$

cannot be zero. For otherwise some $D_j(u_1^{(p)}) = 0$, implying $D_j(u_1) = 0$, contrary to hypothesis.

Now consider the symmetric polynomials

$$P^*(z_1, \ldots, z_m) = \Pi_1^m P(z_i),$$

$$D_j^*(z_1, \ldots, z_m) = \Pi_1^m D_j(z_i).$$

By [9; p. 97], there are polynomials $\pi, \delta_j \epsilon I[z_1, \ldots, z_m]$ such that

$$P^*(z_1, \ldots, z_m) = \pi[\sigma_1(z_1, \ldots, z_m), \ldots, \sigma_m(z_1, \ldots, z_m)],$$

$$D_j^*(z_1, \ldots, z_m) = \delta_j[\sigma_1(z_1, \ldots, z_m), \ldots, \sigma_m(z_1, \ldots, z_m)].$$

It follows from (2.2) that

$$1/p^m = \pi(s_1, \ldots, s_m)/\Pi_1^k[\delta_j(s_1, \ldots, s_m)]^{a(j)}. \tag{2.3}$$

Let $d(j) = \deg[\delta_j]$, and choose integer $b > 0$ so that

$$d = \Sigma_1^k d(j)a(j) + b \geq \deg[\pi];$$

since all $\lambda s_i \epsilon I$, this implies that $\lambda^d \pi(s_1, \ldots, s_m)$ is an integer $w$. Note also that $\lambda^{d(j)}\delta_j(s_1, \ldots, s_m)$ is an integer $w_j$ *independent of* $p$.

Now, (2.3) yields

$$1/p = p^{m-1}w/\lambda^b \Pi_1^k (w_j)^{a(j)}$$

Choosing $p$ to be a prime which divides neither $\lambda$ nor any $w_j$, we obtain a contradiction. (Recall that $\lambda$ does not depend on $p$.)

CASE 2: $N > 1$. Again we assume that $R_N$ contains all the rationals. As in the proof for Case 1, $U = (u_1, \ldots, u_N)$ cannot be algebraically independent over the rationals. Without loss of generality, we assume $u_N$ is algebraically dependent on $\tilde{U} = (u_1, \ldots, u_{N-1})$. Expand the notation to read

$$R_N = R_N(U; D_1, \ldots, D_k).$$

We will show that there exist $\tilde{D}_1, \ldots, \tilde{D}_{k+1} \epsilon \Pi_{N-1}$ with all $\tilde{D}_j(\tilde{U}) \neq 0$ such that $R_{N-1}(\tilde{U}; \tilde{D}_1, \ldots, \tilde{D}_{k+1})$ contains all the rationals. This yields the desired contradiction of the induction hypothesis.

Let $R_0$ be the ring generated over the integers by $\tilde{U}$, i.e.,

$$R_0 = \{\rho : \rho = P(\tilde{U}) \text{ for some } P \epsilon \Pi_{N-1}\},$$

and let $F_0$ be the quotient field of $R_0$. Then $u_N$ is algebraic over $F_0$; let $u_N^{(1)}=u_N$ and let $u_N^{(i)}(2\leq i\leq m)$ be the conjugates of $u_N$ over $F_0$, so that $\Pi_1^m(z-u_N^{(i)})$ is the minimal polynomial of $u_N$ over $F_0$. Then for a suitable $\lambda\epsilon R_0$,

$$g(z)=\lambda\Pi_1^m(z-u_N^{(i)})\epsilon R_0[z].$$

Define $\{\sigma_i\}_1^m$ as above, and $s_i=\sigma_i(u_N^{(1)}, \ldots, u_N^{(m)})$; then all $\lambda s_i\epsilon R_0$.

As in Case 1, we consider a prime $p$. The requirement $1/p\epsilon R_N$ leads as before first to an equation $f(U)=0$ where

$$f\equiv\Pi_1^k(D_j)^{a(j)}-pP,$$

for some appropriate $P\epsilon\Pi_N$ and nonnegative integers $a(j)$. This can be rewritten $f^*(u_N)=0$ where

$$f^*(z)\equiv\Pi_1^k[D_j(\tilde{U},z)]^{a(j)}-pP(\tilde{U},z).$$

Arguing as above we find that all $f^*(u_N^{(i)})=0$, implying

$$1/p=P(\tilde{U},u_N^{(i)})/\Pi_1^k[D_j(\tilde{U},u_N^{(i)})]^{a(j)}$$

and thus

$$1/p^m=\Pi_1^m P(\tilde{U},u_N^{(i)})/\Pi_{i=1}^m\Pi_{j=1}^k[D_j(\tilde{U},u_N^{(i)})]^{a(j)}. \tag{2.4}$$

Now consider the symmetric polynomials

$$P^*(z_1, \ldots, z_m)=\Pi_1^m P(\tilde{U}, z_i)\epsilon R_0[z_1, \ldots, z_m],$$

$$D_j^*(z_1, \ldots, z_m)=\Pi_1^m D_j(\tilde{U}, z_i)\epsilon R_0[z_1, \ldots, z_m].$$

As above, there are polynomials $\pi, \delta_j\epsilon R_0[z_1, \ldots, z_m]$ such that (2.4) yields

$$1/p^m=\pi(s_1, \ldots, s_m)/\Pi_1^k[\delta_j(s_1, \ldots, s_m)]^{a(j)}; \tag{2.5}$$

as before $\lambda$, the $s_i$ and the $\delta_j$ are independent of $p$. Defining $d(j)$, $b$, and $d$ as above, we have

$$1/p=[p^{m-1}\lambda^d\pi(s_1, \ldots, s_m)]/\lambda^b\Pi_1^k(w_j)^{a(j)} \tag{2.6}$$

where $w_j=\lambda^{d(j)}\delta_j(s_1, \ldots, s_m)\epsilon R_0$. We may denote the numerator of (2.6), which lies in $R_0$, as $\tilde{P}(\tilde{U})$. Similarly let $w_j=\tilde{D}_j(\tilde{U})$ for some $\tilde{D}_j\epsilon\Pi_{N-1}$, let $\lambda=\tilde{D}_{k+1}(\tilde{U})$ where $\tilde{D}_{k+1}\epsilon\Pi_{N-1}$, and let $a(k+1)=b$. Then (2.6) yields

$$1/p=\tilde{P}(\tilde{U})/\Pi^{k+1}[\tilde{D}_j(\tilde{U})]^{a(j)}\epsilon R_{N-1}(\tilde{U};\tilde{D}_1, \ldots, \tilde{D}_{k+1}).$$

Thus if $R_N(U; D_1, \ldots, D_k)$ contains all rationals, or equivalently all prime-reciprocals $1/p$, then so does $R_{N-1}(\tilde{U};\tilde{D}_1, \ldots, \tilde{D}_{k+1})$.

This completes the proof.

## 3. The Construction

Our example is constructed as a function of a *denumerable* Markov chain $M$, which will be described next. The state-space of $M$ consists of a *special* state $S_0$, and an infinite lower-triangular array of states $S_{ij}$; the $i$th row of the array consists of the states $S_{i1}, S_{i2}, \ldots, S_{ii}$. The probability of transition from $S_0$ to a row-leading state $S_{i1}$ is $q_i$, where the $q_i$'s are positive numbers (to be speci-

fied) with $\sum_i q_i = 1$; no other states are directly accessible from $S_0$. The other direct transitions are from $S_{i1}$ to $S_{i2}$, $S_{i2}$ to $S_{i3}$, . . ., and from $S_{ii}$ back to $S_0$, each with probability 1.

We now exhibit a vector $\alpha$, with positive components summing to 1, such that

$$\alpha P = \alpha \tag{3.1}$$

where $P$ is the transition matrix of $M$. The entries of $\alpha$ will be denoted $p(S_0)$ and $p(S_{ij})$, with obvious interpretation. For some $x > 0$, set

$$p(S_0) = x, \qquad p(S_{ij}) = q_i x. \tag{3.2}$$

The equations of (3.1) which correspond respectively to $S_0$, to any $S_{i1}$, and to any $S_{i,j+1}$ are

$$\sum_i p(S_{ii}) \cdot 1 = p(S_0),$$

$$p(S_0) q_i = p(S_{i1}),$$

$$p(S_{ij}) \cdot 1 = p(S_{i,j+1}),$$

and all are satisfied; hence (3.1) holds. It only remains to satisfy the normalization

$$p(S_0) + \sum_{j<i} p(S_{ij}) = 1,$$

which identifies $x$ as

$$x = [1 + \sum_i i q_i]^{-1}. \tag{3.3}$$

The $q_i$ will be chosen so that the above makes sense, i.e.,

$$\sum_i i q_i < \infty. \tag{3.4}$$

Take the alphabet $A = \{0, 1\}$. If $\Sigma$ denotes the state-space of $M$, define $f: \Sigma \to A$ by

$$f(S_0) = 0, \quad f(S_{ij}) = 1.$$

$[A, \mu]$ is now defined as the function of $M$ determined by $f$.

Consider the elementary cylinder set

$$C_i = C(\{1,2, \ldots, i+2\}, g_i)$$

where $g_i(1) = g_i(i+2) = 0$ and $g_i = 1$ otherwise. Then $\mu(C_i)$ is the probability of the source producing a pair of 0's separated by exactly $i$ 1's. The sum (1.2) here collapses to a single term, yielding

$$\mu(C_i) = q_i x.$$

Thus *if* $\{q_i\}$ can be selected to include, for each prime $p$, a rational number with denominator divisible by $p$, then the ring generated over the integers by $V(A, \mu) \cup \{1/x\}$ will contain all $q_i$, hence all reciprocals of primes, hence all rationals. By Theorem 1, $[A, \mu]$ cannot be a function of any finite Markov process.

Let $\pi_i$ be the $i$th prime ($\pi_1 = 1$), and set

$$q_i = 1/2^i \pi_i \qquad (i \geq 2). \tag{3.5}$$

Since $\pi_i \geq 2$ in (3.5), we have

$$\sum_{i=2} iq_i < \sum_{i=2} i/2^{i+1} < \infty,$$

and so (3.4) holds for *any* finite $q_1$; since

$$\sum_{i=2} q_i < \sum_{i=2} 1/2^{i+1} = 1/4 < 1,$$

there is a unique $q_1 > 0$ such that $\sum_{i=1} q_i = 1$. Thus the goal described in the previous paragraph has now been accomplished.

Although the main discussion is now complete, certain technical points which were passed over above must still be verified. It must be shown that $M$ is ergodic (as a Markov chain), that the vector $\alpha$ defined in (3.2) is its unique stationary vector, and that $[A, \mu]$ is (as promised) ergodic.

With $\sum$ (as above) the state-space of $M$, the *identity* map of $\sum$ induces an infinite-alphabet source $\left[\sum, \mu^*\right]$ which is a function of $M$. Let $\sum^\infty$ be the set of doubly-infinite sequences $\{\sigma_t\}_{-\infty}^{\infty}$ over $\sum$, and $T^*$ the shift operator on $\sum^\infty$. Then *any* function $[A, \mu]$ of $M$, induced by a function $f$; $\sum \to A$, is ergodic if $\left[\sum, \mu^*\right]$ is ergodic. This follows, in terms of the function $f^\infty : \sum^\infty \to A^\infty$ arising from $f$, from the facts

$$f^\infty T^* = T f^\infty; \qquad \mu(B) = \mu^*[(f^\infty)^{-1}(B)]$$

for any measurable subset $B$ of $A^\infty$.

Thus it suffices to prove that $\left[\sum, \mu^*\right]$ is an ergodic source which, for any $S \epsilon \sum$ and any integer $t$, satisfies

$$\mu^*\{\mathbf{x} \epsilon A^\infty : x_t = S\} = \text{probability of } S \text{ as given by } \alpha. \tag{3.6}$$

The reader is now referred to the discussion in [8; p. 52] of limit properties of Markov chains; we shall use the terminology of that discussion. The chain $M$ is clearly *irreducible* and *nonperiodic*. By (3.4), the mean recurrence time of $S_0$ is finite, so that $S_0$ (and hence all states) are *persistent*. By [8; p. 56], $M$ has a unique invariant instantaneous distribution. Since (3.1) holds, this distribution must be $\alpha$; hence (3.6) holds. In addition, letting $P$ denote the transition matrix of $M$, we have for any $\sigma, \tau \epsilon \sum$

$$\lim_{n \to \infty} (P^n)_{\tau\sigma} = \alpha_\sigma. \tag{3.7}$$

The fact that $\left[\sum, \mu^*\right]$ is ergodic in the sense used in this paper can now be proved from (3.7) by elementary but lengthy arguments; instead we refer the reader to Doob [3; p. 460] for an elegant proof using martingales.

This completes the discussion.

## 4. Conjecture

According to (2.1), *if* $[A, \mu]$ is a function of a finite Markov chain, *then* $V(A, \mu)$ is imbedded in a finately generated ring (of real numbers) over the integers. We conjecture that this necessary condition is *not* sufficient:

CONJECTURE: *There is an information source* $[\mathrm{A}, \mu]$ *on a finite alphabet* A *with the following two properties:*

(1) $\mathrm{V}(\mathrm{A}, \mu)$ *is a finitely generated ring over the integers;*

(2) $[\mathrm{A}, \mu]$ *is not a function of any finite Markov process.*

## 5. References


[1] Dharmadhikari, S. W., Functions of finite Markov chains. Ann. Math. Stat. **34** (1963), 1022–1032.

[2] Dharmadhikari, S. W., Sufficient conditions for a stationary process to be a function of a finite Markov chain, Ann. Math. Stat. **34** (1963), 1033–1041.

[3] Doob, J. L., Stochastic Processes (Wiley, New York, 1953).

[4] Fox, M., and Rubin, H., Functions of processes with Markovian states, Ann. Math. Stat. **39**, 938–946 (1968).

[5] Gilbert, E., On the identifiability problem for finite Markov chains, Ann. Math. Stat. **30**, 688–697 (1959).

[6] Heller, A., On stochastic processes derived from Markov chains, Ann. Math. Stat. **36**, 1286–1291 (1965).

[7] Khinchin, A. I., Mathematical Foundations of Information Theory (Dover, New York, 1957).

[8] Rosenblatt, M., Random Processes (Oxford University Press, New York, 1962).

[9] van der Waerden, B. L., Modern Algebra Vol. I (Ungar, New York, 1953).



(Paper 75B3&4–353)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# On Entire Functions of Exponential Type*

S. M. Shah** and W. C. Sisarcick**

(October 9, 1971)

Let $f$ be an entire function and let $p \geq 1$ and

$$I(l, r)=\left\{\int_0^{2\pi} |f^{(l)}(re^{i\theta})|^p d\theta\right\}^{1/p}.$$

Let $C>0$. If there exists a positive integer $N$ such that for $k=0, 1, \ldots, N$,

$$\sum_{j=0}^{N} \frac{I(k+j, r)}{j!} \geq C \sum_{j=N+1}^{\infty} \frac{I(k+j, r)}{j!},$$

for all sufficiently large $r$, then $f$ is of exponential type not exceeding $\left\{2 \log\left(1+\frac{1}{C}\right)+1+\log (2N)!\right\}$.

If this condition is replaced by related conditions, then also is of exponential type.

Key words: Bounded index; entire function; exponential type; maximum modulus.

## 1. Introduction

An entire function $f(z)$ is said to be of bounded index if and only if there exists a non-negative integer $N$ (independent of $z$) such that

$$\max_{0 \leq j \leq N} \frac{|f^{(j)}(z)|}{j!} \geq \frac{|f^{(k)}(z)|}{k!} \tag{1.1}$$

for all $k$ and all $z$, and the smallest such integer $N$ is called the index of $f(z)$ ([1], [4], [5]).[1] It is known that a function of bounded index $N$ is of exponential type not exceeding $N+1$ [6] but that a function of exponential type need not be of bounded index. In fact any entire function having zeros of arbitrarily large multiplicity is not of bounded index and there exist functions with simple zeros and of exponential type which are not of bounded index [8]. In a recent paper [2] Fred Gross considers interesting variations of condition (1.1) and proves the following

THEOREM A: *Let* f *be entire and* C *a positive constant. If there exists a positive integer* N *such that for* k = 0, 1, . . ., N, f *satisfies one of the following, for all* z *with* | z | *sufficiently large:*

(i) $$\sum_{j=0}^{N} \frac{|f^{(k+j)}(z)|}{j!} > C \sum_{j=N+1}^{\infty} \frac{|f^{(k+j)}(z)|}{j!},$$

(ii) $$\sum_{j=0}^{N} \frac{I(k+j, r)}{j!} > C \sum_{j=N+1}^{\infty} \frac{I(k+j, r)}{j!} \quad (p \text{ some positive integer}),$$

(iii) $$\sum_{j=0}^{N} \frac{M(r, f^{(k+j)})}{j!} > C \sum_{j=N+1}^{\infty} \frac{M(r, f^{(k+j)})}{j!},$$

*then* f *is of exponential type.*

---

*AMS Subject Classification:* Primary 30A64; 30A66.
*An invited paper. Research supported by the National Science Foundation under Grant GP–7544.
**Present address: University of Kentucky, Lexington, Kentucky 40506; Marshall University, Huntington, West Va. 25701.
[1] Figures in brackets indicate the literature references at the end of this paper.

Assuming the known properties of the maximum term and the central index of an entire function we obtain a bound on the exponential type $T$ defined by

$$\limsup_{r\to\infty} \frac{\log M(r,f)}{r} = T.$$

Also we relax the hypothesis in each case. We prove

THEOREM 1: *Let $p \geq 1$ and $C > 0$ be two given constants. Let $f(z)$ be entire, $z = re^{i\Theta}$, and suppose that there exists a positive integer $N$ (independent of $z$) such that for $k = 0,1,2, \ldots, N$, the following inequality*

$$\sum_{j=0}^{N} \frac{I(k+j,r)}{j!} \geq C \sum_{j=N+1}^{\infty} \frac{I(k+j,r)}{j!} \tag{1.2}$$

*holds for all $z$ with $|z|$ sufficiently large. Then $f(z)$ is of exponential type and*

$$T \leq 1 + 2\log\left(1+\frac{1}{C}\right) + \log(2N)!.$$

THEOREM 2: *Let $f(z)$ be entire and $C$ a positive constant. Suppose that there exist two non-negative integers $k$ and $N$ (independent of $z$) such that the following inequality*

$$\sum_{j=0}^{N} \frac{|f^{(k+j)}(z)|}{j!} \geq C \sum_{j=N+1}^{\infty} \frac{|f^{(k+j)}(z)|}{j!} \tag{1.3}$$

*holds for all $z$ with $|z|$ sufficiently large. Then $f(z)$ is of exponential type and*

$$T \leq \max\left\{N, \min_{1\leq j\leq N}\left(\frac{(N+j)!(N+1)}{(N!)C}\right)^{1/j}\right\}.$$

THEOREM 3: *Let $f(z)$ be entire and $C$ a positive constant. Write $M(r, f^{(1)}) = \max|f^{(1)}(z)|$. Suppose that there exist two non-negative integers $k$ and $N$ (independent of $z$) such that the following inequality*

$$\sum_{j=0}^{N} \frac{M(r, f^{(k+j)})}{j!} \geq C \sum_{j=N+1}^{\infty} \frac{M(r, f^{(k+j)})}{j!} \tag{1.4}$$

*holds for all sufficiently large $r$. Then $f(z)$ is of exponential type and*

$$T \leq \left(\frac{(2N+1)!}{C(N!)}\right)^{1/(N+1)}$$

## 2. Proof of Theorem 1.

We have for $r > r_0$

$$\sum_{j=0}^{N} \frac{I(k+j,r)}{j!} = \sum_{j=0}^{N} \frac{I(k+j,r)}{(k+j)!} \frac{(k+j)!}{j!}$$

$$\leq (2N)!\sum_{j=0}^{2N}\frac{I(j,r)}{j!}$$

$$=(2N)!\left\{\left(\sum_{j=0}^{N}+\sum_{j=N+1}^{2N}\right)\frac{I(j,r)}{j!}\right.$$

$$<(2N)!\left(1+\frac{1}{C}\right)\sum_{j=0}^{N}\frac{I(j,r)}{j!}\quad . \tag{2.1}$$

Here we have used the hypothesis with $k=0$ and $r>r_0$ to obtain the last inequality. From the Taylor expansion

$$f^{(k)}(a+h)=\sum_{j=0}^{\infty}\frac{f^{(k+j)}(a)}{j!}h^j$$

we have

$$|f^{(k)}(re^{i\theta})|^p=\left|\sum_{j=0}^{\infty}\frac{f^{(k+j)}((r-1)e^{i\theta})}{j!}e^{ij\theta}\right|^p$$

and so

$$I(k,r)\leq\left\{\int_0^{2\pi}\left\{\sum_{j=0}^{\infty}\frac{|f^{(k+j)}((r-1)e^{i\theta})|}{j!}\right\}^p d\theta\right\}^{1/p}\quad .$$

We now use Minkowski inequality [3, p. 148] to obtain

$$I(k,r)\leq\sum_{j=0}^{\infty}\frac{I(k+j,r-1)}{j!}.$$

Using the hypothesis and (2.1), we have for $r>r_1>1+r_0$,

$$I(k,r)\leq\left(1+\frac{1}{C}\right)\sum_{j=0}^{N}\frac{I(k+j,r-1)}{j!}\quad .$$

$$\leq\left(1+\frac{1}{C}\right)^2(2N)!\sum_{j=0}^{N}\frac{I(j,r-1)}{j!}\quad .$$

Hence

$$\sum_{k=0}^{N}\frac{I(k,r)}{k!}\leq e\left(1+\frac{1}{C}\right)^2(2N)!\sum_{j=0}^{N}\frac{I(j,r-1)}{j!}.$$

Write $\lambda=e\left(1+\frac{1}{C}\right)^2(2N)!$ and $\sum_{k=0}^{N}\frac{I(k,r)}{k!}=\xi(r)$. Then $\xi(r)\leq\lambda\,\xi(r-1)$ and so we get, for $r>r_1$,

$$\xi(r)<C_1\lambda^r, \tag{2.2}$$

where $C_1 = C_1(N, p, r_1)$ is a constant. Write now $f(z) = \sum_0^\infty a_n z^n$, $\mu(r) = \max_n |a_n| r^n$. Then

$$|a_n| r^n \leq \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})| \, d\theta$$

$$\leq \frac{1}{2\pi} \left\{ \int_0^{2\pi} |f(re^{i\theta})|^p d\theta \right\}^{1/p} \left\{ \int_0^{2\pi} d\theta \right\}^{1/q}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Hence (2.2) gives

$$\mu(r) < (2\pi)^{-1/p} C_1 \lambda^r. \tag{2.3}$$

Now [10; 32–34]

$$\limsup_{r \to \infty} \frac{\log \log \mu(r)}{\log r} = \limsup_{r \to \infty} \frac{\log \log M(r)}{\log r} = \rho \text{ (say)},$$

and if $\rho < \infty$, then

$$\log \mu(r) \sim \log M(r).$$

These two relations along with (2.3) imply that $f(z)$ is of exponential type and $T \leq \log \lambda$. The proof is complete.

## 3. Proof of Theorem 2

The proof of this theorem is essentially the same as that of Theorem 1 of [6]. We therefore outline the proof. Write $f^{(k)}(z) = F(z)$. Then

$$\sum_{j=0}^{N} \frac{|F^{(j)}(z)|}{j!} \geq C \sum_{j=N+1}^{\infty} \frac{|F^{(j)}(z)|}{j!}.$$

Hence for $p \geq 1$

$$\sum_{j=0}^{N} \frac{M(r, F^{(j)})}{j!} \geq C \frac{M(r, F^{(N+p)})}{(N+p)!}.$$

Let

$$\limsup_{r \to \infty} \frac{\nu(r, f)}{r} = a.$$

Then (cf: [6]) there exists a sequence $r_n \uparrow \infty$ such that for $r = r_n$, $(n > n_0)$,

$$\frac{M(r, F^{(j)})}{M(r, F^{(q)})} \sim \left( \frac{\nu(r, F)}{r} \right)^{j-q} \geq b^{j-q}, \qquad 0 \leq q \leq j, \; j = 1, 2, \ldots, 2N, \; b < a.$$

Hence if $b_1 < b$, we get for $1 \leq j \leq N$, *and* $r = r_n$,

$$CM(r, F^{(N+j)}) < (N+j)! \left\{ \frac{M(r, F^{(N+j)})}{b_1^{N+j}} + \ldots + \frac{M(r, F^{(N+j)})}{N!\, b_1^j} \right\},$$

and so we have for $1 \leq j \leq N$,

$$\frac{Cb_1^{N+j}}{(N+j)!} \leq 1 + \frac{b_1}{1!} + \ldots \frac{b_1^N}{N!}. \tag{3.1}$$

If $b_1 < N$ then we get $a \leq N$ and so $T \leq N$. If $b_1 \geq N$ then we have from (3.1)

$$\frac{Cb_1^{N+j}}{(N+j)!} \leq (N+1)\frac{b_1^N}{N!}$$

and so

$$T \leq \min_{1 \leq j \leq N} \left\{ \frac{(N+j)!(N+1)}{N!C} \right\}^{1/j}.$$

The proof is complete.

The proof of Theorem 3 is similar and omitted.

## 4. Remarks and Examples

(a) There are functions satisfying all hypotheses (i)–(iii) of Theorem A. Take, for instance, $f(z) = e^{az}$, $a > 0$. Gross [2] has proved that a periodic function of exponential type satisfies (ii) and (iii) of Theorem A for sufficiently large $r$ and $i = 0, 1, \ldots, N$.

(b) Let $f(z)$ be entire and suppose that it satisfies the differential equation

$$P_0(z) f^{(k)}(z) + P_1(z) f^{(k-1)}(z) + \ldots + P_k(z) f(z) = Q(z), \tag{4.1}$$

where $P_j(z)$, $j = 0, 1, \ldots, k$, $Q(z)$ are polynomials and $P_0(z)$ ($\not\equiv 0$) is of degree not less than that of any $P_j(z)$. Then $f(z)$ is of bounded index [7]. Furthermore $f(z)$ satisfies the hypothesis of Theorem 2, that is, inequality (1.3) with $k = 0$ and for all $z$. For let $\alpha = 1 + C$ and $F(z) = f(\alpha z)$. Then $F(z)$ also satisfies a differential equation of the form (4.1). Hence $F(z)$ is of bounded index $N$ say, and for all $k$ and all $z$

$$\frac{|F^{(k)}(z)|}{k!} \leq \max_{0 \leq s \leq N} \frac{|F^{(s)}(z)|}{s!}.$$

This implies that

$$\frac{|f^{(k)}(z)|}{k!} \leq \alpha^{N-k} \max_{0 \leq s \leq N} \frac{|f^{(s)}(z)|}{s!}.$$

Hence for $k = N+1, N+2, \ldots,$

$$\frac{M(r, f^{(k)})}{k!} \leq \alpha^{N-k} \Omega(r)$$

where

$$\Omega(r) = \max_{0 \leq s \leq N} \left\{ \frac{M(r, f^{(s)})}{s!} \right\}.$$



449-764 OL - 72 - 4

This gives

$$\sum_{k=N+1}^{\infty} \frac{M(r,f^{(k)})}{k!} \leq \frac{1}{C}\Omega(r) < \frac{1}{C}\sum_{0}^{N} \frac{M(r,f^{(s)})}{s!}.$$

Similarly the entire solution $f(z)$ of (4.1) satisfies the hypothesis of Theorem 3 with $k=0$ and for all $z$. It also satisfies the hypothesis of Theorem 1, that is, given $p \geq 1$ and $C > 0$, there exists a positive integer $N$ such that for $k=0, 1, 2, \ldots, N$ and for all $r=|z|$, (1.2) holds. To prove this we let $\alpha=1+C$ and denote by $M \equiv M_0, M_1, \ldots, M_M$, the indices of the functions $f(\alpha z) \equiv F(z)$, $F'(z), \ldots, F^{(M)}(z)$ and let

$$N = \max_{0 \leq J \leq M} M_J.$$

Then the inequality (1.2) holds for $k=0, 1, \ldots, N$. We omit the details.

(c) We now show, in Example 1, that there exist entire functions which satisfy the hypothesis of Theorem 3 (with $k=0$, $N=0$ and given $C$) but which do not satisfy the hypothesis of Theorem 2. In Example 2 we give a function which does not satisfy the hypothesis (i) of Theorem A but which satisfies the hypothesis of corresponding Theorem 2.

EXAMPLE 1: Let $C > 0$ be a given constant. Let

$$f(z) = \prod_{n=1}^{\infty} \left(1 + \frac{z}{na^{n+1}}\right)^n$$

where we choose $a \geq 2$ such that $\dfrac{1}{a(a-1)} \leq \log(1+C)$. Then $f(z)$ is entire and satisfies the conditions of Theorem 3, with $k=0$, $N=0$ and given $C$. Since $f$ has zeros of arbitrarily large multiplicity, the hypothesis of Theorem 2 cannot be satisfied.

EXAMPLE 2: Let $a_1 = k_1 = 10$; $k_{j+1} = k_j^2$,

$$a_{j+1} = k_{j+1} \exp\left\{\left(\log \frac{3}{2}\right)k_j\right\}, \; j \geq 1$$

$$f(z) = \prod_{j=1}^{\infty} \left(1 - \frac{z}{a_j}\right)^{k_j}, \; F(z) = f(z) - 1.$$

Then $F(z)$ is an entire function of bounded index [9] and $F'(z)$ has zeros of arbitrarily large multiplicity. Let $C > 0$, $a = 1+C$, $g(z) = F\left(\dfrac{z}{a}\right)$. Since $F(z)$ has an infinity of zeros, the index $N$ of $F(z)$ is a positive integer and we have for all $z$ and $k$,

$$\max_{0 \leq j \leq N} \frac{|F^{(j)}(z)|}{j!} \geq \frac{|F^{(k)}(z)|}{k!}.$$

Hence for all $z$

$$\sum_{j=N+1}^{\infty} \frac{|g^{(j)}(z)|}{j!} = \sum_{j=N+1}^{\infty} \frac{|F^{(j)}(z/a)|}{a^j j!}$$

$$\leqslant \sum_{j=N+1}^{\infty} \frac{1}{a^j} \max_{0\leqslant k\leqslant N} \frac{|F^{(k)}(z/a)|}{k!}$$

$$= \frac{a}{a^{N+1}(a-1)} \max_{0\leqslant k\leqslant N} \frac{a^k |g^k(z)|}{k!}$$

$$< \frac{a^{N+1}}{a^{N+1}(a-1)} \sum_{k=0}^{N} \frac{|g^{(k)}(z)|}{k!}$$

$$= \frac{1}{C} \sum_{k=0}^{N} \frac{|g^{(k)}(z)|}{k!}.$$

This shows that $g(z)$ satisfies the hypothesis of Theorem 2 with $k=0$. Further $g'(z)$ has zeros of arbitrarily large multiplicity and so $g(z)$ cannot satisfy the hypothesis (*i*) (when $k=1$) of Theorem A.

## 5. References


[1] Gross, F., Entire functions of bounded index, Proc. Amer. Math. Soc. **18** (1967), 974–980.

[2] Gross, F., Entire functions of exponential type, J. Res. Nat. Bur. Stand. (U.S.), 72B (Math. Sci.), No. 1, 55–59 (Jan.–Feb. 1970).

[3] Hardy, G. H., Littlewood, J. E., Polya, G., Inequalities (Camb. Univ. press 1964).

[4] Lepson, B., Differential equations of infinite order, hyperdirichlet series and entire functions of bounded index, Lecture Notes. Summer Institute on Entire Functions, Univ. of California, La Jolla (1966).

[5] Lepson, B., Differential equations of infinite order, hyperdirichlet series and entire functions of bounded index, Proc. Symp. in Pure Math. **XI,** 298–308 (1968).

[6] Shah, S. M., Entire functions of bounded index, Proc. Amer. Math. Soc. **19,** 1017–1022 (1968).

[7] Shah, S. M., Entire functions satisfying a linear differential equation, J. of Math. and Mech. **18,** 131–136 (1968).

[8] Shah, S. M., Entire functions of unbounded index and having simple zeros, Math. Zeit., **118,** 193–196 (1970).

[9] Shah, S. M., On entire functions of bounded index whose derivatives are of unbounded index, Jour. Lond. Math. Soc. (2) 4, 127–139 (1971).

[10] Valiron, G., Lectures on the general theory of integral functions (Chelsea, New York 1949).


(Paper 75B3&4–354)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# A Table of Integrals of the Error Function. II. Additions and Corrections*

**Murray Geller** and Edward W. Ng****

**(September 28, 1971)**

This is an extension of a compendium of indefinite and definite integrals of products of the error function with elementary or transcendental functions recently published by the authors.

Key words: Error functions; indefinite integrals; special functions.

## 1. Introduction

Since the present authors published an extensive compendium of integrals involving the error function [1],[1] numerous comments and suggestions have been received. In particular, we have been advised by P. I. Hadji (П. И. ХАДЖИ) of his publication in Russian [2], which contains formulas not included in [1]. Careful examination of [2] has revealed numerous mathematical and typographical errors. Bearing in mind also the inaccessibility of his report, the authors believe it to be appropriate to publish an extension of [1] based largely on corrected versions of Hadji's formulas but also containing new results. It should be noted that the authors have verified and/or corrected all of Hadji's formulas in addition to rederiving and checking all of the formulas in [1].

Throughout this article, we conform to the format and notation of [1], including section and equation numbers. A supplementary glossary is also included. Within each subsection, typographical errors, equivalent forms and errors detected in [1] will be listed under corrections.

## 2. Supplementary Glossary

| | | |
|---|---|---|
| $B(p, q)$ | Beta Function | $\Gamma(p)\Gamma(q)/\Gamma(p+q)$ |
| $\beta(x)$ | | $\frac{1}{2}[\psi(\frac{1}{2}+\frac{1}{2}x)-\psi(\frac{1}{2}x)]$ |
| $C_n^{\lambda}(x)$ | Gegenbauer Polynomial | |
| $G$ | Catalan's Constant | 0.915965594 . . . |
| $P_n^{(\alpha,\beta)}(x)$ | Jacobi Polynomial | |
| $T_n(x)$ | Chebyshev Polynomial | |

*AMS Subject Classification:* Primary 6505.
* An invited paper. This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract #NAS7-100, sponsored by the National Aeronautics and Space Administration.

** Present address: Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103.

[1] Figures in brackets indicate the literature references at the end of this paper.

## 3.* Integrals of Products of Error Functions with Other Functions

### 3.1. Combination of Error Function with Powers

*Additions*

(20) $$\int_a^\infty \operatorname{erf}(x)x^{-p}dx=\frac{1}{p-1}\left[\frac{\operatorname{erf}(a)}{a^{p-1}}+\frac{1}{\sqrt{\pi}}\frac{e^{-a^2/2}}{a^{p/2}}W_{-p/4,(p-2)/4}(a^2)\right], \qquad p>1$$

(21) $$\int_0^\infty [\operatorname{erf}(ax+c)-\operatorname{erf}(bx+c)]\frac{dx}{x}=ln\left(\frac{a}{b}\right)\operatorname{erfc}(c)$$

(22) $$\int_0^\infty [b\operatorname{erf}(ax)-a\operatorname{erf}(bx)]\frac{dx}{x^2}=\frac{2ab}{\sqrt{\pi}}ln\left(\frac{b}{a}\right)$$

(23) $$\int_0^\infty [\sqrt{x}\operatorname{erf}(\sqrt{x})-\operatorname{erf}(x)]\frac{dx}{x^2}=\frac{1}{\sqrt{\pi}}(2-\gamma)$$

(24) $$\int_0^\infty \left[\frac{\operatorname{erf}(x)}{x}-\frac{2}{\sqrt{\pi}}\frac{1}{1+x^2}\right]\frac{dx}{x}=\frac{1}{\sqrt{\pi}}(2-\gamma)$$

(25) $$\int_0^u (u-x)^p\operatorname{erf}[\sqrt{(a/x)}]dx=\frac{u^{p+1}}{p+1}\left[1-\frac{\Gamma(p+2)}{\sqrt{\pi}}\left(\frac{u}{a}\right)^{1/4}e^{-a/(2u)}W_{-p-(5/4),1/4}\left(\frac{a}{u}\right)\right], \qquad p>-1$$

(26) $$\int_0^\infty [(b+ix)^{-p-1}-(b-ix)^{-p-1}]\operatorname{erf}(ax)dx=-\frac{2i}{p}(a\sqrt{2})^pe^{a^2b^2/2}D_{-p}(ab\sqrt{2}), \qquad p>0$$

(27) $$\int_0^\infty x[(b+ix)^{-p-1}+(b-ix)^{-p-1}]\operatorname{erf}(ax)dx$$

$$=\frac{2}{p(p-1)}(\alpha/b)^{p-1}e^{a^2b^2/2}[(p-1)\alpha D_{-p}(\alpha)-pD_{1-p}(\alpha)], \qquad p>1,\ \alpha=ab\sqrt{2}$$

(28) $$\int_{-\infty}^\infty \frac{\operatorname{erf}(ax)}{(x+b)^2}dx=2a\sqrt{\pi}e^{-a^2b^2}\operatorname{erfi}(ab)$$

(29) $$\int_0^\infty \frac{\operatorname{erfc}[a(1+x)^{1/4}]}{(1+x)^{3/4}}dx=\frac{4}{a\sqrt{\pi}}e^{-a^2}-4\operatorname{erfc}(a)$$

(30) $$\int \operatorname{erf}[\sqrt{(ax)}]\frac{x-b}{\sqrt{x}(x+b)^2}dx=-\frac{2\sqrt{x}}{x+b}\operatorname{erf}[\sqrt{(ax)}]-2\sqrt{(a/\pi)}e^{ab}E_1(ab+ax)$$

(31) $$\int_0^u x^{2\nu-1}(u^2-x^2)^{\mu-1}\operatorname{erf}(ax)dx$$

$$=\frac{a}{\sqrt{\pi}}u^{2\mu+2\nu-1}B(\mu,\nu+\tfrac{1}{2})_2F_2(\tfrac{1}{2},\nu+\tfrac{1}{2};\frac{3}{2},\mu+\nu+\tfrac{1}{2};-a^2u^2),\ \mu>0,\ \nu>-\tfrac{1}{2}$$

(32) $$\int_0^\infty \frac{x\operatorname{erf}(ax)}{(x^2+b^2)^2}dx=\frac{a}{2b}\sqrt{\pi}\,e^{a^2b^2}\operatorname{erfc}(ab)$$

*Section 3 corresponds to section 4 of reference [1].

(33) $$\int_0^1 \frac{x \operatorname{erf}(ax)}{(x^2+1)^2}\,dx = \frac{a}{4}\sqrt{\pi}\,e^{a^2}\,[1-\operatorname{erf}^2(a)] - \tfrac{1}{4}\operatorname{erf}(a)$$

(34) $$\int_1^\infty \frac{x \operatorname{erf}(ax)}{(x^2+1)^2}\,dx = \frac{a}{4}\sqrt{\pi}\,e^{a^2}[1-\operatorname{erf}(a)]^2 + \tfrac{1}{4}\operatorname{erf}(a)$$

(35) $$\int_0^\infty \frac{x \operatorname{erf}(ax)}{(x^2+b^2)^{p+1}}\,dx = \frac{a}{2}\,b^{1-2p}\,\frac{\Gamma(p-\frac{1}{2})}{\Gamma(p+1)}\,{}_1F_1(\tfrac{1}{2};\tfrac{3}{2}-p;a^2b^2)$$

$$+\frac{a^{2p}}{2\sqrt{\pi}}\,\frac{\Gamma(\frac{1}{2}-p)}{\Gamma(p+1)}\,{}_1F_1(p;p+\tfrac{1}{2};a^2b^2),\quad p>0,\ p\neq n+\tfrac{1}{2}$$

(36) $$\int_0^\infty \frac{\operatorname{erf}(ax)}{(x^2+b^2)^{3/2}}\,dx = \frac{1}{b^2}\,[1-\exp(a^2b^2)\operatorname{erf} c\,(ab)]$$

(37) $$\int_0^\infty \frac{x \operatorname{erf}(ax)}{(x^2+b^2)^{3/2}}\,dx = \frac{a}{\sqrt{\pi}}\exp(a^2b^2/2)K_0\left(\frac{a^2b^2}{2}\right)$$

(38) $$\int_0^\infty \frac{x \operatorname{erf}[a\sqrt{(x^2+b^2)}]}{(x^2+b^2)^{3/2}}\,dx = \frac{1}{b}\operatorname{erf}(ab) + \frac{a}{\sqrt{\pi}}\,E_1(a^2b^2)$$

(39) $$\int_0^\infty \frac{x \operatorname{erf}[\sqrt{(ax)}]}{(x^2+4b^2)^{3/2}}\,dx = \frac{a\pi}{2\sqrt{2}}\,[J_{1/4}(ab)Y_{-1/4}(ab) - J_{-1/4}(ab)Y_{1/4}(ab)]$$

### 3.2. Combination of Error Function with Exponentials and Powers

*Corrections*

(10) Replace $(bz^n + nz^{n-1})$ in last expression by $(bz^{n-1} + (n-1)z^{n-2})$.

(12) Replace $(bz^n + nz^{n-1})$ in last expression by $\dfrac{1}{a\sqrt{\pi}}\,(bz^{n-1} + (n-1)z^{n-2})$.

(13) Replace $(2a)^{(k-n)/2}$ by $(2a^2)^{(k-n)/2}$.

*Additions*

(15) $$\int_0^\infty [\operatorname{erf}(e^{-ax}) - \operatorname{erf}(e^{-bx})]\,\frac{dx}{x} = \operatorname{erf}(1)\ln\left(\frac{b}{a}\right)$$

(16) $$\int_{-\infty}^\infty \operatorname{erf}\left[\frac{ae^{-ipx}}{(b\pm ix)^\alpha}\right]\frac{dx}{x^2+c^2} = \frac{\pi}{c}\operatorname{erf}\left[\frac{ae^{\mp pc}}{(b\pm c)^\alpha}\right],\ p>0,\ \alpha\geq 0,\ b\neq c \text{ for the lower sign}$$

(17) $$\int_{-\infty}^\infty \operatorname{erf}\left(\frac{a}{b\pm ix}\right)e^{-ipx}\,\frac{dx}{x^2+c^2} = \frac{\pi}{c}\,e^{\mp pc}\operatorname{erf}\left(\frac{a}{b\pm c}\right),\ p>0,\ b\neq c \text{ for the lower sign}$$

(18) $$\int_{-\infty}^\infty (ix)^{-p}\operatorname{erf}(ae^{-ibx})\,\frac{dx}{x^2+c^2} = \frac{\pi}{c^{p+1}}\operatorname{erf}(ae^{-bc}),\ |p|<1$$

(19) $$\int_{-\infty}^\infty \operatorname{erf}(ae^{\pm ibx})\,\frac{dx}{x-c} = \pm i\pi \operatorname{erf}(ae^{\pm ibc})$$

(20) $$\int_0^\infty \frac{dx}{\sin h\frac{1}{2}\pi x}\left[\frac{\operatorname{erf}(1+ix)}{(1+ix)^2}-\frac{\operatorname{erf}(1-ix)}{(1-ix)^2}\right]=2i\ \operatorname{erf}(1)$$

### 3.3. Combination of Error Function with Exponentials of More Complicated Arguments

*Corrections*

(2) Equivalent form: $\displaystyle\int_0^\infty \operatorname{erf}(ax)e^{-b^2x^2}dx=\frac{1}{b\sqrt{\pi}}\tan^{-1}\left(\frac{a}{b}\right)$

(3) Remove condition "$b$ may be complex" and replace by $\mathscr{R}(a^2)>\mathscr{R}(b^2)$

(8) For the form given, add the condition $\mathscr{R}(b^2)>\mathscr{R}(a^2)$

Equivalent form:

$$\int_0^\infty \operatorname{erf}(ax)e^{-b^2x^2}x^p dx=\frac{a}{b^{p+1}\sqrt{\pi}}\frac{\Gamma(1+\frac{1}{2}p)}{(a^2+b^2)^{\frac{1}{2}}}\,{}_2F_1\left(\tfrac{1}{2}-\tfrac{1}{2}p,\ \tfrac{1}{2};\ \tfrac{3}{2};\ \frac{a^2}{a^2+b^2}\right),$$

$$\mathscr{R}(a^2)>0,\ \mathscr{R}(b^2)>0,\ p>-2.$$

(16) $$\int_1^\infty \operatorname{erf} c\,(ax)e^{a^2x}x^{-3}dx=\frac{1}{2}\,(1+2a^2)e^{a^2}\operatorname{erf} c\,(a)-a\pi^{-1/2}$$

(35) Add the condition $\mathscr{R}(c^2)>\mathscr{R}(a^2)$

(36) Replace $K_b(a^2)$ by $K_b(\frac{1}{2}a^2)$

*Additions*

(39) $$\int x\exp(x^2)\operatorname{erf}(x)dx=\frac{1}{2}\exp(x^2)\operatorname{erf}(x)-\frac{1}{\sqrt{\pi}}x$$

(40) $$\int_{-a}^{a}\exp(x^2)\operatorname{erf} c(x)dx=\sqrt{\pi}\operatorname{erf} i\,(a)$$

(41) $$\int[\operatorname{erf}(ax)]^n\exp(-a^2x^2)dx=\frac{\sqrt{\pi}}{2a(n+1)}[\operatorname{erf}(ax)]^{n+1}$$

(42) $$\int[a\operatorname{erf}(bx)\exp(-a^2x^2)+b\operatorname{erf}(ax)\exp(-b^2x^2)]dx=\frac{\sqrt{\pi}}{2}\operatorname{erf}(ax)\operatorname{erf}(bx)$$

(43) $$2b^2\int\operatorname{erf}(ax+c)\exp(-b^2x^2)xdx$$

$$=\frac{a}{\sqrt{(a^2+b^2)}}\operatorname{erf}\left(x\sqrt{(a^2+b^2)}+\frac{ac}{\sqrt{(a^2+b^2)}}\right)\exp\left(-\frac{b^2c^2}{a^2+b^2}\right)-\operatorname{erf}(ax+c)\exp(-b^2x^2)$$

(44) $$\int x^{2n}\operatorname{erf}(ax)\exp(-a^2x^2)\,dx=\frac{(2n)!}{n!}\frac{\sqrt{\pi}}{(2a)^{2n+1}}A,$$

where

$$A=\frac{1}{2}\operatorname{erf}^2(ax)-\frac{2a}{\sqrt{\pi}}x\operatorname{erf}(ax)\exp(-a^2x^2)\sum_{k=0}^{n-1}\frac{k!}{(2k+1)!}(4a^2x^2)^k$$

$$-\frac{\exp(-2a^2x^2)}{\pi}\sum_{k=0}^{n-1}2^k\frac{k!\,k!}{(2k+1)!}e_k(2a^2x^2)$$

(45) $$\int x^{2n+1}\operatorname{erf}(ax)\exp(-a^2x^2)\,dx=\frac{n!}{a^{2n+2}}B,$$

where

$$B=2^{-3/2}\operatorname{erf}(ax\sqrt{2})\sum_{k=0}^{n}\binom{2k}{k}8^{-k}-\frac{1}{2}\operatorname{erf}(ax)e^{-a^2x^2}e_n(a^2x^2)$$

$$-\frac{ax}{\sqrt{\pi}}e^{-2a^2x^2}\sum_{k=1}^{n}\binom{2k}{k}8^{-k}\sum_{l=0}^{k-1}\frac{l!}{(2l+1)!}(8a^2x^2)^l$$

(46) $$2b^2\int x^{2n+1}\operatorname{erf}(ax)\exp(-b^2x^2)\,dx=2n\int x^{2n-1}\operatorname{erf}(ax)\exp(-b^2x^2)\,dx$$

$$-x^{2n}\operatorname{erf}(ax)\exp(-b^2x^2)+\frac{2a}{\sqrt{\pi}}\int x^{2n}\exp[-(a^2+b^2)x^2]\,dx$$

(47) $$\int\operatorname{erf}(ax)\exp(-a^2x^2)\frac{dx}{x^2}=-\frac{1}{x}\operatorname{erf}(ax)\exp(-a^2x^2)-\frac{a\sqrt{\pi}}{2}\operatorname{erf}^2(ax)-\frac{a}{\sqrt{\pi}}E_1(2a^2x^2)$$

(48) $$\int\exp\{-b^2[\operatorname{erf}(ax)]^p\}\exp(-a^2x^2)[\operatorname{erf}(ax)]^q dx$$

$$=\frac{\sqrt{\pi}}{2ap}b^{-2(1+q)/p}\gamma\left(\frac{q+1}{p},\,b^2[\operatorname{erf}(ax)]^p\right),\quad q>-1,\ p>0$$

(49) $$\int\frac{e^{-a^2x^2}}{\operatorname{erf}(ax)}dx=\frac{\sqrt{\pi}}{2a}\ln[\operatorname{erf}(ax)]$$

(50) $$\int_c^\infty\operatorname{erf}(ax)e^{-b^2x^2}\sin hcx\,dx=\frac{\sqrt{\pi}}{2b}\exp[c^2/(4b^2)]\operatorname{erf}\left[\frac{ac}{2b\sqrt{(a^2+b^2)}}\right]$$

(51) $$\int_0^\infty\operatorname{erf}\left[a\sqrt{(x^2+b^2)}\right]e^{-c^2x^2}x\,dx=\frac{1}{2c^2}\operatorname{erf}(ab)+\frac{a}{2c^2}\frac{e^{b^2c^2}}{\sqrt{(a^2+c^2)}}\operatorname{erf}c\left[b\sqrt{(a^2+c^2)}\right]$$

(52) $$\int_0^\infty\operatorname{erf}\left[\sqrt{(ax)}\right]e^{-b/x}\frac{dx}{x^2}=\frac{1}{b}\left[1-e^{-2\sqrt{(ab)}}\right]$$

(53) $$\int_{-\infty}^\infty\operatorname{erf}(ae^{x/2})e^{x/2}\exp(-e^x)dx=\frac{2}{\sqrt{\pi}}\tan^{-1}a$$

(54) $$\int_0^\infty\left[\frac{a\operatorname{erf}(\alpha\exp[-ce^{ax}])}{1-e^{-ax}}-\frac{b\operatorname{erf}(\alpha\exp[-ce^{bx}])}{1-e^{bx}}\right]dx=\ln\left(\frac{b}{a}\right)\operatorname{erf}(\alpha e^{-c})$$

(55) $$\int_0^\infty\operatorname{erf}i(ax)e^{-b^2x^2}\frac{dx}{x}=\sin^{-1}\left(\frac{a}{b}\right),\ b\geq a$$

(56) $$\int_a^\infty e^{x^2}(x^2-a^2)^{-1/2}\operatorname{erfc}(x)\,dx=\frac{1}{2}e^{a^2/2}K_0(a^2/2)$$

(57) $$\int_0^\infty \operatorname{erf} i(ax)e^{-a^2x^2}\frac{x}{x^2+b^2}\,dx=\frac{\pi}{2}\operatorname{erfc}(ab)e^{a^2b^2}$$

### 3.4. Definite Integrals from Laplace Transforms Involving Erf $(\sqrt{ax})$

*Corrections*

(12) Replace erf $(\sqrt{ax})$ in integrand by erfc $(\sqrt{ax})$ and remove the condition $b>a$.

(13) First term inside square brackets in integrand should read $(2ax+1)xe^{ax}\operatorname{erfc}(\sqrt{ax})$ instead of $(2a^{1/2}x^{1/2}+1)xe^{ax}\operatorname{erf}(\sqrt{ax})$ and remove the condition $b>a$.

(14) Last term inside square brackets in integrand should read

$$-2(ax^3/\pi)^{1/2}(2ax+5) \text{ instead of } -2(a^3x^5/\pi)^{1/2}(2ax+5).$$

(15) Replace erfc $(\sqrt{cx})$ in integrand by erf $(\sqrt{cx})$ and add the condition $b>c$.

(16) First term inside square brackets in integrand should read

$$a^{1/2}e^{cx}\operatorname{erf}(\sqrt{cx}) \text{ instead of } a^{1/2}e^{cx}\operatorname{erf}(\sqrt{bx}).$$

(18) First term inside brackets in integrand should read

$$a(x/\pi)^{1/2}e^{-a^2/(4x)}.$$

*Additions*

(19) $$\int_0^\infty e^{-x}\frac{\operatorname{erf}(\sqrt{x})}{(1+e^{-x})^2}\,dx=1+(\sqrt{2}-1)\zeta\left(\frac{1}{2}\right)$$

(20) $$\int_0^\infty e^{-x}\operatorname{erf}(\sqrt{ax})\left[\frac{1-e^{-nx}}{1-e^{-x}}\right]\frac{dx}{\sqrt{x}}=\frac{2}{\sqrt{\pi}}\sum_{k=1}^{n}\frac{1}{\sqrt{k}}\tan^{-1}\left[\sqrt{(a/k)}\right]$$

(21) $$\int_0^\infty e^{-bx}\operatorname{erf}(\sqrt{ax})\,e^{-ncx}[\sin hcx]^n\frac{dx}{\sqrt{x}}$$
$$=\frac{2}{\sqrt{\pi}}\frac{1}{2^n}\sum_{k=0}^{n}(-1)^k\binom{n}{k}(b+2ck)^{-1/2}\tan^{-1}\left(\frac{a}{b+2ck}\right)^{1/2}$$

### 3.5. Combination of Error Function with Trigonometric Functions

*Corrections*

(10) For the form given $a<b$.

For $a>b$:

$$\int_0^\infty \operatorname{erf}(ax)\sin b^2x^2dx=\frac{1}{4b\sqrt{(2\pi)}}\left[\ln\left(\frac{a^2+b^2+ab\sqrt{2}}{a^2+b^2-ab\sqrt{2}}\right)+2\tan^{-1}\left(\frac{ab\sqrt{2}}{b^2-a^2}\right)+2\pi\right]$$

(11) Change the condition on $p$ to $\mathscr{R}(p) > -2$.

(12) Change the condition on $p$ to $\mathscr{R}(p) > -1$.

(13) $$\int_0^\infty \operatorname{erf}(ax) \sin bx \frac{dx}{x^2} = \frac{b}{2} E_1\left(\frac{b^2}{4a^2}\right) + a\sqrt{\pi} \operatorname{erf}\left(\frac{b}{2a}\right).$$

(16) $$\int_0^\infty \operatorname{erf}\left(\sqrt{\frac{a}{x}}\right) \sin bx dx = \frac{1}{b}\left[1 - e^{-A} \cos A\right], \qquad A = (2ab)^{1/2}$$

(17) $$\int_0^\infty \operatorname{erf}\left(\sqrt{\frac{a}{x}}\right) \cos bx dx = \frac{1}{b} e^{-A} \sin A, \qquad A = (2ab)^{1/2}$$

(24) $$\int_0^\infty \operatorname{erf} c(ax) x \sin x \sin hx dx = \frac{1}{2}\left[\frac{1}{a^2} \sin\left(\frac{1}{2a^2}\right) + \cos\left(\frac{1}{2a^2}\right) - 1\right]$$

*Additions*

(31) $$2i(b^2 + c^2) \int \operatorname{erf}(ax) e^{-bx} \sin cx dx = (b + ic)A_- - (b - ic)A_+ ,$$

where

$$A_\pm = \exp\left[\left(\frac{b \pm ic}{2a}\right)^2\right] \operatorname{erf}\left(ax + \frac{b \pm ic}{2a}\right) - \exp\left[-(b \pm ic)x\right] \operatorname{erf}(ax)$$

(32) $$\int \tan[\operatorname{erf}(ax)] e^{-a^2x^2} dx = -\frac{\sqrt{\pi}}{2a} \ln\{\cos[\operatorname{erf}(ax)]\}$$

(33) $$\int_0^\infty \operatorname{erf}(ax) e^{-b^2x^2} \sin cx dx = \frac{\sqrt{\pi}}{2b} e^{-c^2/(4b^2)} \operatorname{erf} i\left[\frac{ac}{2b(a^2 + b^2)^{1/2}}\right]$$

(34) $$\int_0^\pi \operatorname{erf}(a \sin x) \cos(2n+1)x \, dx = \int_0^\pi \operatorname{erf}(a \sin x) \sin 2nx \, dx = 0$$

(35) $$\int_0^{2\pi} \operatorname{erf}(a \sin x + b \cos x) dx = 0$$

(36) $$\int_0^{2\pi} \operatorname{erf}(ae^{ix}) \sin nx \, dx = \begin{cases} 0 & n = 2k \\ 2\sqrt{\pi} \dfrac{(-1)^k}{k!} \dfrac{a^{2k+1}}{2k+1} i & n = 2k+1 \end{cases}$$

(37) $$\int_0^{2\pi} \operatorname{erf}(ae^{ix}) \cos nx \, dx = \begin{cases} 0 & n = 2k \\ 2\sqrt{\pi} \dfrac{(-1)^k}{k!} \dfrac{a^{2k+1}}{2k+1} & n = 2k+1 \end{cases}$$

(38) $$\int_0^{\pi/2} \operatorname{erf}(a \tan x) \sin 2x \, dx = \sqrt{\pi} \, ae^{a^2} \operatorname{erf} c(a)$$

(39) $$\int_0^{\pi/2} \operatorname{erf} c(a \tan x) \cos 2x \, dx = \pi^{-1/2} ae^{a^2} E_1(a^2)$$

(40) $$\int_0^1 \operatorname{erf}(ax)\cos[(2n+1)\cos^{-1}x](1-x^2)^{-1/2}dx$$
$$=(-1)^n\frac{\Gamma\left(n+\frac{1}{2}\right)}{(2n+1)!}\frac{e^{-a^2/2}}{2a}M_{-1/2,\,n+1/2}(a^2)$$

(41) $$a\int_0^\infty \operatorname{erf} c[b^{1/2}\{(x^2+c^2)^{1/2}-c\}^{1/2}]\sin axdx=1-\frac{1}{\Delta}\left[\frac{1}{2}b(b+\Delta)\right]^{1/2}e^{-c(\Delta-b)},\quad \Delta=(a^2+b^2)^{1/2}$$

(42) $$\int_0^\infty [\operatorname{erf}(\tan^{-1}ax)-\operatorname{erf}(\tan^{-1}bx)]\frac{dx}{x}=\ln\left(\frac{a}{b}\right)\operatorname{erf}\left(\frac{\pi}{2}\right)$$

### 3.6. Combination of Error Function With Logarithms and Powers

*Corrections*

(5) Replace $\frac{1}{2}z^2 \operatorname{erf}(az)\ln z$ on the R.H.S. by $\frac{1}{2}z^2 \operatorname{erf} c(az)\ln z$.

*Additions*

(10) $$\int_0^\infty \operatorname{erf}(ax)\ln(2\pm 2\cos x)\frac{x}{x^2+c^2}dx=-\pi\ln(1\pm e^{-c})\operatorname{erf} i(ac)$$

(11) $$\int_0^1 \operatorname{erf}(a\ln x)\frac{x^{b-1}}{1+x^c}dx=\frac{1}{c}\beta\left(\frac{b}{c}\right)\operatorname{erf}\left(\frac{a}{2c}\right)$$

(12) $$\int_0^\infty \operatorname{erf}(a\ln x)\frac{x^{b-1}}{1+x^c}dx=\frac{1}{c}\left[\beta\left(\frac{b}{c}\right)-\beta\left(1-\frac{b}{c}\right)\right]\operatorname{erf}\left(\frac{a}{2c}\right),\qquad c>b$$

(13) $$\int_0^\infty \ln\left[\frac{c+d\operatorname{erf}(ax)}{c+d\operatorname{erf}(bx)}\right]\frac{dx}{x}=\ln\left(\frac{a}{b}\right)\ln\left(\frac{c+d}{c}\right)$$

(14) $$\int_0^{\pi/2} \operatorname{erf}(b\sin 2x)\ln(a\tan x)dx=\frac{1}{2}(\ln a)\operatorname{erf}(2b)\frac{[\Gamma(\frac{1}{2}a)]^2}{\Gamma(a)}$$

(15) $$\int_0^{\pi/2} \operatorname{erf}(a\ln\tan x)\,dx=0$$

(16) $$\int_0^\infty \operatorname{erf}\left(\frac{ax}{x^2+b^2}\right)\ln\left(\frac{x}{b}\right)\frac{dx}{x}=0$$

(17) $$\int_0^\infty \operatorname{erf}\left(\frac{x^p}{x^{2p}+b^{2p}}\right)\ln\left(\frac{x}{b}\right)\frac{dx}{x^2+b^2}=0$$

(18) $$\int_0^\infty \operatorname{erf}(x^p+x^{-p})\ln x\frac{dx}{x}=0$$

(19) $$\int_0^\infty \operatorname{erf}(x^p+x^{-p})\ln x\frac{dx}{1+x^2}=0$$

(20) $$\int_0^\infty \frac{\operatorname{erf}(a \ln x)}{\ln x} \frac{dx}{1-x^2}=0$$

(21) $$\int_0^\infty \operatorname{erf}(a \ln x) \frac{dx}{1+bx+x^2}=0, \quad |b|<2$$

(22) $$2a^2 \int_0^\infty x \ln x \operatorname{erf} c(x) e^{-a^2x^2}dx$$

$$=-\frac{1}{2}\gamma-\ln[1+(1+a^2)^{1/2}]+\frac{1}{2}(a^2+1)^{-1/2}[\gamma+2\ln 2+\ln(1+a^2)]$$

### 3.7. Combination of Two Error Functions

*Corrections*

(3) Last term on R.H.S. should read $+\frac{2b}{\sqrt{\pi}} E_1(2ab)$

(4) Last term on R.H.S. should read $-\frac{2b}{\sqrt{\pi}} E_1(2ab)$

(7) $$\int_0^1 \operatorname{erf}(x) \operatorname{erf}[(1-x^2)^{1/2}]x dx=(2e)^{-1}$$

(8) Remove the entire expression and replace by

$$12b^3\sqrt{\pi}\int_0^\infty x^2 \operatorname{erf} c(bx) \operatorname{erf} c\left(\frac{a}{2x}\right) dx=e^{-ab}(4+ab-a^2b^2)+a^3b^3E_1(ab).$$

(9) Add the condition $c>a$.

*Additions*

(10) $$\int \operatorname{erf}(ax) \operatorname{erf}(bx)dx=x \operatorname{erf}(ax) \operatorname{erf}(bx)-\frac{1}{ab\sqrt{\pi}}(a^2+b^2)^{1/2} \operatorname{erf}[x\sqrt{(a^2+b^2)}]$$

$$+\frac{1}{a\sqrt{\pi}} \operatorname{erf}(bx)e^{-a^2x^2}+\frac{1}{b\sqrt{\pi}} \operatorname{erf}(ax)e^{-b^2x^2}$$

(11) $$(2n+1)\int x^{2n} \operatorname{erf}(ax) \operatorname{erf}(bx)dx$$

$$=x^{2n+1} \operatorname{erf}(ax) \operatorname{erf}(bx)-\frac{2b}{\sqrt{\pi}}\int x^{2n+1} \operatorname{erf}(ax)e^{-b^2x^2}dx-\frac{2a}{\sqrt{\pi}}\int x^{2n+1} \operatorname{erf}(bx)e^{-a^2x^2}dx$$

(12) $$\int x^{2n} \operatorname{erf}^2(ax)dx=\frac{x^{2n+1}}{2n+1} \operatorname{erf}^2(ax)-\frac{4}{2n+1}\frac{n!}{\sqrt{\pi}\,a^{2n+1}}A,$$

where

$$A=\frac{1}{2\sqrt{2}}\operatorname{erf}(ax\sqrt{2})\sum_{k=0}^{n}\binom{2k}{k}8^{-k}-\frac{1}{2}\operatorname{erf}(ax)e^{-a^2x^2}e_n(a^2x^2)$$

$$-\frac{ax}{\sqrt{\pi}}e^{-2a^2x^2}\sum_{k=0}^{n}\binom{2k}{k}8^{-k}\sum_{l=0}^{k-1}\frac{l!}{(2l+1)!}(8a^2x^2)^l$$

(13) $$\int x\operatorname{erf}^2(ax)dx=\left(\frac{x^2}{2}-\frac{1}{4a^2}\right)\operatorname{erf}^2(ax)+\frac{x}{a\sqrt{\pi}}\operatorname{erf}(ax)e^{-a^2x^2}+\frac{1}{2\pi a^2}e^{-2a^2x^2}$$

(14) $$\int x^{2n+1}\operatorname{erf}^2(ax)dx=\left[x^{2n+2}-\frac{2}{n!}\frac{(2n+1)!}{(2a)^{2n+2}}\right]\frac{\operatorname{erf}^2(ax)}{2n+2}+\frac{2}{(2a)^{2n+2}}\frac{(2n+1)!}{(n+1)!}B,$$

where

$$B=\frac{2ax}{\sqrt{\pi}}\operatorname{erf}(ax)e^{-a^2x^2}\sum_{k=0}^{n}\frac{k!}{(2k+1)!}(4a^2x^2)^k+\frac{e^{-2a^2x^2}}{\pi}\sum_{k=0}^{n}\frac{2^k k!k!}{(2k+1)!}e_k(2a^2x^2)$$

(15) $$\int_0^\infty \operatorname{erf}(ax)\operatorname{erf}(bx)\frac{dx}{x^2}=\frac{2}{\sqrt{\pi}}\left[a\ln\left(\frac{b+c}{a}\right)+b\ln\left(\frac{a+c}{b}\right)\right],\ c=(a^2+b^2)^{1/2}$$

(16) $$\int_0^\infty x^{p-1}\operatorname{erf}c(ax)\operatorname{erf}(bx)dx=\frac{2b}{a^{p+1}}\frac{\Gamma(1+p/2)}{\pi(p+1)}{}_3F_2\left(\frac{1}{2},\frac{p+1}{2},\frac{p+2}{2};\frac{3}{2},\frac{p+3}{2};-\frac{b^2}{a^2}\right),$$

$$a>b,\qquad p>-1$$

(17) $$\int_0^\infty \operatorname{erf}(ax)\operatorname{erf}(bx)e^{-c^2x^2}\frac{dx}{x^2}$$

$$=\frac{2}{\sqrt{\pi}}\left\{a\ln\left[\frac{b+\Delta}{(a^2+c^2)^{1/2}}\right]+b\ln\left[\frac{a+\Delta}{(b^2+c^2)^{1/2}}\right]-c\tan^{-1}\left(\frac{ab}{c\Delta}\right)\right\},$$

$$\Delta=(a^2+b^2+c^2)^{1/2}$$

(18) $$\int_0^\infty \operatorname{erf}(ax)\operatorname{erf}(bx)e^{-c^2x^2}dx=\frac{1}{c\sqrt{\pi}}\tan^{-1}\left(\frac{ab}{c\Delta}\right),\ \Delta=(a^2+b^2+c^2)^{1/2}$$

(19) $$\int_0^\infty \operatorname{erf}(ax)\operatorname{erf}(bx)e^{-c^2x^2}xdx=\frac{1}{c^2\pi}\left[\frac{b}{\Delta_2}\tan^{-1}\frac{a}{\Delta_2}+\frac{a}{\Delta_1}\tan^{-1}\frac{b}{\Delta_1}\right]$$

$$\Delta_1=(a^2+c^2)^{1/2},\ \Delta_2=(b^2+c^2)^{1/2}$$

(20) $$\int_0^\infty xe^{x^2}\ln x\operatorname{erf}c^2(x)dx=\frac{2}{\pi}\left[G-1-\frac{\gamma}{2}\left(1-\frac{\pi}{4}\right)\right]$$

(21) $$\int_0^\infty [1-\operatorname{erf}^3(x)]dx=\frac{6}{\pi}\left(\frac{2}{\pi}\right)^{1/2}\tan^{-1}(2^{-1/2})$$

(22) $$\int_0^\infty [1-\operatorname{erf}^4(x)]dx=\frac{12}{\pi}\left(\frac{2}{\pi}\right)^{1/2}\tan^{-1}(8^{-1/2})$$

*Corrections*

(12) Replace $\Gamma\left(-p-1, \frac{b^2}{4a^2}\right)$ on the R.H.S. by $\Gamma\left(-p-\frac{1}{2}, \frac{b^2}{4a^2}\right)$.

(13) Replace $W_{-p/2,\,p/2}\left(\frac{b^2}{8a^2}\right)$ by $W_{-p/2,\,p/2}\left(\frac{b^2}{4a^2}\right)$;

replace $b^{p+1/2}/a^{3/2p+1}$ by $b^p/a^{2p+1}$;

change the condition $-1 < p < \frac{3}{2}$ to $-\frac{1}{2} < p < \frac{3}{4}$.

(19) Multiply the R.H.S. by $\frac{2}{\sqrt{\pi}}$;

change the condition on $p$ to $-\frac{1}{2} < p < 0$.

(20) Multiply the R.H.S. by $\frac{2}{\sqrt{\pi}}$

(21) Replace the condition $\lambda + \frac{1}{2}p > 0$ by $\lambda + \frac{1}{2}p > -\frac{1}{2}$

*Additions*

(22) $$\int_0^\infty J_1[a \operatorname{erf}(x)]e^{-x^2}dx = \frac{\sqrt{\pi}}{2a}[1 - J_0(a)]$$

(23) $$\int_0^\infty J_0[a\sqrt{\{1-\operatorname{erf}^2(x)\}}]e^{-x^2}dx = \frac{\sqrt{\pi}}{2a}\sin a$$

(24) $$\int_0^\infty K_0(ax)\operatorname{erf}(bx)x\,dx = \frac{1}{8b^2}e^c[K_1(c) - K_0(c)],\ c = a^2/(8b^2)$$

(25) $$\int_0^\infty e^{-x^2}I_0(x^2)\operatorname{erf} c(ax)dx = (8\pi)^{-1/2}\ln\left[1 + \frac{2}{a^2} + \frac{2}{a^2}(1+a^2)^{1/2}\right]$$

(26) $$\int_0^\infty x^{2p-2}e^{-x^2}K_\nu(x^2)\operatorname{erf}(x)dx = \frac{\Gamma(p+\nu)\Gamma(p-\nu)}{2^p\Gamma\left(p+\frac{1}{2}\right)}\,{}_3F_2\left(p+\nu, p-\nu, \frac{1}{2}; p+\frac{1}{2}, \frac{3}{2}; -\frac{1}{2}\right),$$

$$p > |\nu|$$

(27) $$\int_0^\infty x^{2\lambda}J_{2\nu}\left(\frac{a}{x}\right)\operatorname{erf}(bx)\,dx$$

$$= \left(\frac{a}{2}\right)^{2\nu}\frac{b^{2\nu-2\lambda-1}\Gamma(\lambda-\nu+1)}{\sqrt{\pi}(2\nu-2\lambda-1)\Gamma(2\nu+1)}\,{}_1F_3\left(\nu-\lambda-\frac{1}{2}; 2\nu+1, \nu-\lambda, \nu-\lambda+\frac{1}{2}; \frac{a^2b^2}{4}\right)$$

$$+\left(\frac{a}{2}\right)^{2\lambda+2}\frac{b\Gamma(\nu-\lambda-1)}{\sqrt{\pi}\,\Gamma(\nu+\lambda+2)}\,{}_1F_3\left(\frac{1}{2}; \frac{3}{2}, \lambda-\nu+2, \lambda+\nu+2; \frac{a^2b^2}{4}\right),$$

$$\nu > \lambda > -\frac{5}{4}$$

*Corrections*

(3) Replace $\mathrm{si}(2p\,x)$ in integrand by $\frac{1}{a}\,\mathrm{si}\left(\frac{x}{a}\right)$.

(5) Replace $e^{a^2/4}$ on R.H.S. by $e^{a^2/2}$;

change second index of $W$ from $(1+2\mu)/4$ to $(1+2\nu)/4$;

change argument of $W$ from $a^2$ to $a^2/2$;

change the condition $\mu < \nu$ to $\mu > \nu$.

(6) Add the condition $\nu > -\frac{1}{2}$.

(7) Add the condition $\nu > -\frac{1}{2}$.

(9) Change the R.H.S. to

$$\frac{1}{\sqrt{\pi}}\frac{\Gamma(b)}{p^{2b-1}}\,(2a-2b+1)^{-1}\,{}_2F_1(a, a-b+\tfrac{1}{2}; a-b+\tfrac{3}{2}; -1), \qquad a+\tfrac{1}{2} > b$$

(10) Change the R.H.S. to

$$\frac{1}{\sqrt{\pi}}\frac{\Gamma(b)}{p^{2b-1}}\,(p/q)^{2a}(2a-2b+1)^{-1}\,{}_2F_1\left(a, a-b+\tfrac{1}{2}; a-b+\tfrac{3}{2}; -\frac{p^2}{q^2}\right), \qquad a+\tfrac{1}{2} > b;\ q^2 \geq p^2$$

(11) Change the R.H.S. to

$$\frac{1}{\sqrt{\pi}}\frac{p}{q^{\nu+2}}\Gamma\left(1+\frac{\nu}{2}\right)\frac{\Gamma(b)\Gamma(a-1-\frac{1}{2}\nu)}{\Gamma(a)\Gamma(b-1-\frac{1}{2}\nu)}\,{}_3F_2\left(\tfrac{1}{2}, 1+\frac{\nu}{2}, 2-b+\frac{\nu}{2}; \tfrac{3}{2}, 2-a+\frac{\nu}{2}; -\frac{p^2}{q^2}\right), \qquad \nu > -2;\ q^2 \geq p^2$$

(16) Change the condition on $p$ to $p > 0$.

(18) Integrand should contain $\mathrm{erfc}\left(\sqrt{\frac{a}{2}}\,x\right)$ instead of $\mathrm{erfc}\left(\frac{a}{\sqrt{2}}\,x\right)$;
coefficient on R.H.S. should read

$$2^{\nu/2+1}a^{-3/2}\Big/\left[\left(1+\frac{\nu}{2}\right)\Gamma\left(1-\frac{\nu}{2}\right)\right] \text{ instead of } 2^{\nu/2+2}a^{-3/2}/(\nu\pi)$$

(20) Replace $p^{\nu+1/2}$ on the R.H.S. by $a^{\nu+1/2}$

(21) First term inside ${}_3F_2$ expression should read

$$-\lambda+\mu+\tfrac{1}{2} \text{ instead of } \lambda+\mu+\tfrac{1}{2};$$

argument of ${}_3F_2$ should read $+a$ instead of $-a$;

add the condition $a \leq 1$; $a=1$, $\lambda > \frac{1}{2}p$.

(22) $$\int_0^1 T_{2n+1}(x)\ \text{erf}\ (ax)(1-x^2)^{-1/2}dx=(-1)^n\frac{\Gamma(n+\frac{1}{2})}{2a(2n+1)!}e^{-a^2/2}M_{-1/2,\ n+1/2}(a^2)$$

(23) $$\int_{-\infty}^{\infty}\{T_n[\text{erf}\ (x)]\}^2e^{-x^2}dx=\sqrt{\pi}\left(\frac{2n^2-1}{4n^2-1}\right)$$

(24) $$\int_0^1 P_n(1-2x^2)\ \text{erf}\ (ax)dx$$

$$=\left(\frac{a}{2}\right)^{2n+1}\frac{1}{(2n+1)\Gamma(n+\frac{3}{2})}\,{}_2F_2(n+\tfrac{1}{2},n+1;n+\tfrac{3}{2},2n+2;-a^2)$$

(25) $$\int_0^1 P_{2n+1}(x)\ \text{erf}\ (ax)\ \frac{dx}{\sqrt{x}}=(-1)^{n+1}\frac{1}{2a\sqrt{(a\pi)}}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(2n+\frac{5}{2})}\,e^{-a^2/2}M_{-3/4,\ n+3/4}(a^2)$$

(26) $$\int_0^1(1-x^2)\ P_{2n+1}^{(p,p)}(x)\ \text{erf}\ (ax)dx$$

$$=\frac{(-1)^n}{2\sqrt{\pi}}\frac{\Gamma(n+\frac{1}{2})}{(2n+1)!}\frac{\Gamma(2n+p+2)}{\Gamma(2n+p+\frac{5}{2})}\frac{e^{-a^2/2}}{a^{p+3/2}}\,M_{-p/2-3/4,\ n+p/2+3/4}(a^2),\quad p>-1$$

(27) $$\int_0^\infty x^{\lambda-1}(1-x^2)^{-\mu/2}\,p_\nu^\mu(x)\ \text{erf}\ (ax)dx$$

$$=\frac{2^{\mu-\lambda}a\Gamma(\lambda+1)}{\Gamma(1+\frac{1}{2}\lambda-\frac{1}{2}\mu-\frac{1}{2}\nu)\Gamma(\frac{3}{2}+\frac{1}{2}\lambda-\frac{1}{2}\mu+\frac{1}{2}\nu)}\,{}_3F_3\left(\tfrac{1}{2},\frac{\lambda+1}{2},\frac{\lambda+2}{2};\tfrac{3}{2},\frac{2+\lambda-\mu-\nu}{2},\frac{3+\lambda-\mu-\nu}{2};-a^2\right),$$

$$\lambda>-1;\ \mu<1$$

(28) $$\int\text{erf}\ (ax)H_n(ax)dx=[2a(n+1)]^{-1}\left\{\text{erf}\ (ax)H_{n+1}(ax)+\frac{2}{\sqrt{\pi}}H_n(ax)e^{-a^2x^2}\right\}$$

(29) $$\int_0^\infty e^{-x^2}H_{2n+1}(x)\ \text{erf}\ (x)dx=(-1)^n\frac{(2n)!}{2^{n+1/2}n!}$$

(30) $$\int_0^\infty e^{-x^2}H_{2n}(x)\ \text{erf}\ (x)xdx=(-1)^{n+1}\frac{(2n)!}{2^{n+3/2}n!}\frac{2n+1}{2n-1}$$

(31) $$\int_0^\infty e^{-x^2}H_n(x)H_{n+2m+1}(x)\ \text{erf}\ (x)dx=(-1)^m\frac{(2m+2n+2)!}{(m+n+1)!}\frac{1}{2^{m+n+3/2}}\frac{1}{2m+1}$$

(32) $$\int_0^\infty x^{p-1}H_{2n}(ax)\ \text{erf}\,c\,(bx)dx=(-1)^n\frac{(2n)!\,\Gamma(\frac{1}{2}+\frac{1}{2}p)}{n!\quad pb^p\sqrt{\pi}}\,{}_3F_2\left(-n,\frac{p}{2},\frac{p+1}{2};\frac{1}{2},\frac{p+2}{2};\frac{a^2}{b^2}\right),$$

$$p>0;\ b^2>a^2$$



449-764 OL - 72 - 5

(33) $$\int_0^1 (1-x^2)^{p-1/2}\, C_{2n+1}^{p}(x)\ \text{erf}\ (ax)\, dx$$

$$= \frac{(-1)^n}{2^{2p}} \frac{a^{2n+1}}{(2n+1)!} \frac{\Gamma(n+\frac{1}{2})}{\Gamma(p)} \frac{\Gamma(2n+2p+1)}{\Gamma(2n+p+2)}\ {}_1F_1(n+\tfrac{1}{2};\ 2n+p+2;\ -a^2)$$

$$p>0$$

(34) $$\int_0^\infty e^{-a^2x^2/4} D_{2n+1}(ax)\ \text{erf}\ (bx)dx = (-1)^n \frac{(2n-1)!!}{a} \left(\frac{2b^2}{a^2+2b^2}\right)^{n+1/2}$$

(35) $$\int_0^\infty e^{-x^2/4}\,[D_{2p}(x) - D_{2p}(-x)]\ \text{erf}\ (x)dx = \left(\frac{4}{3}\right)^p \sqrt{(2/\pi)}\ \sin p\pi\ \Gamma(p),\ p>0$$

(36) $$\int_0^\infty x^{2p-1} e^{-x^2/4} D_{2q}(x)\ \text{erf}\ (\tfrac{1}{2}x)dx$$

$$= 2^{q-p-1/2} \frac{\Gamma(2p+1)}{\Gamma(p-q+1)}\ {}_3F_2\ (\tfrac{1}{2},\ p+\tfrac{1}{2},\ p+1;\ \tfrac{3}{2},\ p-q+1;\ -\tfrac{1}{2}),\ p>-\tfrac{1}{2}$$

(37) $$\int_0^\infty x^{2n+1} e^{-x^2/2} \Gamma_n^{n+1/2}\ (x^2/2)\ \text{erf}\ (x)dx$$

$$= \frac{1}{\sqrt{\pi}} \left(\frac{2}{3}\right)^{n+1/2} \frac{2^n}{n!} \frac{\Gamma(2n+3/2)}{(n+\frac{1}{2})}\ {}_2F_1\ (-n,\ n+\tfrac{1}{2};\ n+\tfrac{3}{2};\ \tfrac{2}{3})$$

(38) $$\int_0^\infty L_n^a(a^2x^2)\,\text{erf}\ c(bx)x^{p-1}dx$$

$$= \binom{\alpha+n}{n} \frac{\Gamma(\frac{1}{2}+\frac{1}{2}p)}{pb^p\sqrt{\pi}}\ {}_3F_2\left(-n,\ \frac{p}{2},\ \frac{p+1}{2};\ \alpha+1,\ \frac{p+2}{2};\ \frac{a^2}{b^2}\right),\ \alpha>-1;\ p>0;\ b^2>a^2$$

(39) $$\int_{-\infty}^{\infty} \frac{[\text{erf}\ (ae^{i\pi x}) - \text{erf}\ (ae^{-i\pi x})]}{\Gamma(\alpha+x)\Gamma(\beta-x)\ \sin \pi x} dx = i \frac{2^{\alpha+\beta-1}}{\Gamma(\alpha+\beta-1)}\ \text{erf}\ (a),\ \alpha+\beta>1$$

(40) $$\int_0^\infty x^{p-1}\ \text{erf}\ c(bx)\gamma(\alpha,\ a^2x^2)dx$$

$$= \frac{a^{2\alpha}\Gamma(\alpha+\frac{1}{2}+\frac{1}{2}p)}{\alpha\sqrt{\pi}\,(p+2\alpha)b^{p+2\alpha}}\ {}_3F_2\left(\alpha,\ \alpha+\tfrac{1}{2}p,\ \alpha+\tfrac{1}{2}+\tfrac{1}{2}p;\ \alpha+1,\ \alpha+1+\tfrac{1}{2}p;\ -\frac{a^2}{b^2}\right),\quad p+2\alpha>0;\ b^2>a^2$$

(41) $$\int_0^\infty \text{erf}\ (ax)\,{}_2F_1(\alpha+1,\ \beta+1+3/2;\ -c^2x^2)xdx$$

$$= \frac{1}{4\alpha\beta a^2} \left(\frac{a}{c}\right)^{\alpha+\beta+1} e^{a^2/(2c^2)} W_{(1-\alpha-\beta)/2,\,(\alpha-\beta)/2}\left(\frac{a^2}{c^2}\right),\ \alpha>0;\ \beta>0$$

(42) $$\int_0^\infty x^{4p}e^{-a^2x^2}{}_1F_1\left(\frac{1}{2}-2p;\ 2p+1;\ a^2x^2\right)\ \text{erf}\ (bx)dx$$

$$=\frac{1}{2a\sqrt{\pi}}\,\Gamma(2p)\left[\frac{b^2}{a^2(a^2+b^2)}\right]^{2p}{}_2F_1\left(\frac{1}{2}-2p,\ 2p;\ 2p+1;\ \frac{b^2}{a^2+b^2}\right),$$

$$p>0$$

(43) $$\int_0^\infty e^{x^2}x^{2\beta-2}{}_1F_1\ (\alpha;\ \beta;-\gamma^2x^2)\ \text{erf}\ c(x)dx$$

$$=\frac{\Gamma(2\beta-1)\Gamma(1+\alpha-\beta)}{2^{2\beta-1}\Gamma(\frac{1}{2}+\alpha)}\,{}_2F_1\ (\alpha,\ \beta-\tfrac{1}{2};\ \alpha+\tfrac{1}{2};\ 1-\gamma^2),$$

$$\gamma^2<2;\ 1+\alpha>\beta>\tfrac{1}{2}$$

(44) $$\int_0^\infty e^{-x^2}\Psi(\alpha,\ \beta;\ \gamma^2x^2)\ \text{erf}\ i(x)xdx$$

$$=\frac{1}{2\gamma^3}\,\frac{\Gamma(\alpha-\frac{1}{2})\Gamma(5/2-\beta)}{\Gamma(\alpha)\Gamma(\alpha-\beta+2)}\,{}_2F_1\left(1,\frac{5}{2}-\beta;\ \alpha-\beta+2;\ 1-\frac{1}{\gamma^2}\right),$$

$$\alpha>\tfrac{1}{2};\ \gamma^2>\tfrac{1}{2};\ \beta<\tfrac{5}{2}$$

(45) $$\int_0^\infty e^{-x/2}x^{p-1}W_{\alpha,\beta}(x)\ \text{erf}\ (a^{1/2}x^{1/2})\,dx$$

$$=2\sqrt{(a/\pi)}\,\frac{\Gamma(1+\beta+p)\Gamma(1-\beta+p)}{\Gamma(3/2-\alpha+p)}\,{}_3F_2\ (\tfrac{1}{2},\ 1+\beta+p,\ 1-\beta+p;\ \tfrac{3}{2},\ \tfrac{3}{2}-\alpha+p;\ -a),$$

$$p+1>|\beta|$$

## 4. References


[1] Ng, Edward W., and Geller, Murray, A Table of Integrals of the Error Functions, J. Res. Nat. Bur. Stand. (U.S.), 73B (Math. Sci.), No. 1, 1–20 (Jan.–Mar. 1969).

[2] Hadji, P. I., The Calculation of Some Integrals Containing a Probability Function, Moldavian Academy of Sciences SSR; Series of Physical, Tech. and Math. Sci. **2**, 81 (1968).



(Paper 75B3&4–355)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December, 1971

# A Note on the Time Dependence of the Effective Axis and Angle of a Rotation*

**Harry Gelman****

(October 19, 1971)

The time dependent rotation of one orthogonal coordinate system with respect to a fixed one is considered in the parametrization based on the effective axis and angle of the rotation, a parametrization which has recently been used to discuss the irreducible representations of the rotation group. The method of the intrinsic vector is used to derive the equations of motion for the instantaneous effective axis and angle. A new representation of the angular velocity is obtained in a rotating orthogonal coordinate system generated by a unit vector along the effective axis, and a new geometrical interpretation of the effective angle is given.

Key words: Angular velocity; effective axis and angle of rotation; intrinsic vector; kinematics of a rigid body; rotation group; spatial rotation.

## 1. Introduction

The most common parametrization of the rotation group is by means of the Euler angles [1, 2].[1] Recently, there has been some interest in the parametrization of this group by the direct use of the effective axis and angle of a rotation. Moses [3, 4] has calculated the irreducible representations in that parametrization, and [5] has computed the orthogonality relations between the matrices of any two such irreducible representations. Carmeli [6] has obtained the same results using a technique which differs from Moses', and which is originally due to Weyl. The outstanding feature of these results is their remarkable simplicity, compared to the corresponding results in the parametrization through the Eulerian angles.

The most familiar use of the Euler angles is of course in the description of the rotational motion of an orthogonal coordinate system. The instantaneous angular velocity is related to the time rate of change of the orientation of the rotating coordinate system by the well known equations [7]

$$\omega_1 = \frac{d\theta}{dt}\cos\phi + \frac{d\psi}{dt}\sin\theta\sin\phi \tag{1a}$$

$$\omega_2 = \frac{d\theta}{dt}\sin\phi - \frac{d\psi}{dt}\sin\theta\cos\phi \tag{1b}$$

$$\omega_3 = \frac{d\phi}{dt} + \frac{d\psi}{dt}\cos\theta, \tag{1c}$$

where $\omega_1$, $\omega_2$, and $\omega_3$ are the components of the angular velocity in a space-fixed coordinate system, and $\phi$, $\theta$, and $\psi$ are the Euler angles which describe the instantaneous orientation of the rotating coordinate system in the spaced-fixed system. If $\omega(t)$ is regarded as known, eqs (1) are the differential equations governing the time development of the orientation of the rotating system.

---

*AMS Subject Classification:* Primary 2080.

* An invited paper.

** Present address: Department D–81, The MITRE Corporation, Bedford, Mass. 01730.

[1] Figures in brackets indicate the literature references at the end of this paper.

With the increased use of the effective axis and angle to parametrize a rotation, it is of interest to examine the relation between the instantaneous angular velocity of a rotating coordinate system and the time development of its orientation when the orientation is specified by an effective axis and angle. Such an examination would give additional insight into this simple, yet rarely used parametrization of rotations. That examination is the purpose of this paper.

In section 2, the method of the intrinsic vector [8, 9] is used to derive expressions for the time rate of change of the effective axis and angle of rotation, induced by the angular velocity. A decomposition is then given of the angular velocity in an orthogonal set of vectors generated by a unit vector along the effective axis. This decomposition is the analog of eqs (1) in the parametrization based on the effective axis and angle. It is used to construct a new geometrical interpretation of the effective angle of rotation. The question of the extent to which the motion of the effective axis differs from the rigid rotation of an axis fixed in the rotating system, is briefly considered. It is shown that the motion of the effective axis fulfills the condition of rigid rotation only in the simple case when the rotating coordinate system rotates about a fixed axis.

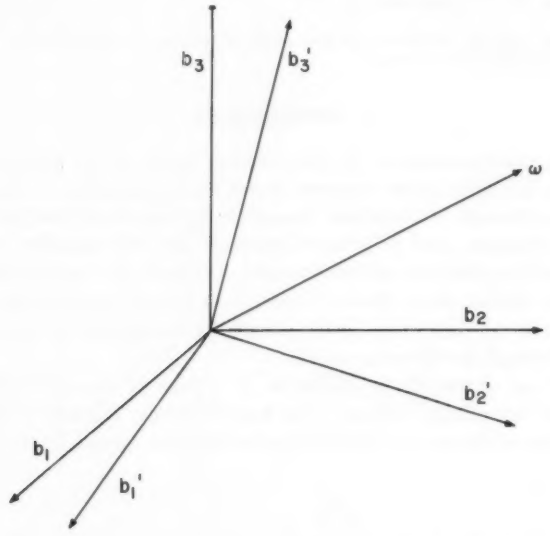


FIGURE 1. *Instantaneous orientation of a coordinate system rotating with respect to a fixed one.*

## 2. Time Development of the Orientation of a Rotating Coordinate System

We consider a pair of co-original orthogonal coordinate systems $S$ and $S'$, of the same handedness, specified, respectively, by the triads of unit vectors $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ and $(\mathbf{b}_1', \mathbf{b}_2', \mathbf{b}_3')$, and take the point of view that $S'$ is rotating with respect to $S$. The situation is shown in figure 1, in which $\boldsymbol{\omega}$ represents the instantaneous angular velocity of $S'$. At any instant $t$ the triads $\{\mathbf{b}_i\}$ and $\{\mathbf{b}_i'\}$ are linearly related by the orthogonal rotation dyadic $\mathbf{A}(t)$ according to the equation

$$\mathbf{b}_i'(t) = \mathbf{b}_i \cdot \mathbf{A}(t) = \tilde{\mathbf{A}}(t) \cdot \mathbf{b}_i\,, \tag{2}$$

where $\tilde{\mathbf{A}}$ is the dyadic conjugate to $\mathbf{A}$. The matrix elements of $\mathbf{A}$ in the coordinate system $S$ are given by

$$A_{ij} = \mathbf{b}_i \cdot \mathbf{A} \cdot \mathbf{b}_j = \mathbf{b}_i' \cdot \mathbf{b}_j. \tag{3}$$

The time derivative of $\tilde{\mathbf{A}}$ and the instantaneous angular velocity are related by

$$\frac{d\tilde{\mathbf{A}}}{dt} = \boldsymbol{\omega} \times \tilde{\mathbf{A}} = (\boldsymbol{\omega} \times \mathbf{I}) \cdot \tilde{\mathbf{A}}, \tag{4}$$

in which $\mathbf{I}$ is the unit dyadic. Using the fact that the conjugate of $\boldsymbol{\omega} \times \mathbf{I}$ is $-\boldsymbol{\omega} \times \mathbf{I}$, and the orthogonality of $\mathbf{A}$,

$$\tilde{\mathbf{A}} \cdot \mathbf{A} = \mathbf{A} \cdot \tilde{\mathbf{A}} = \mathbf{I}, \tag{5}$$

we may write eq (4) as

$$\frac{d\mathbf{A}}{dt} = -\mathbf{A} \cdot (\boldsymbol{\omega} \times \mathbf{I}). \tag{6}$$

In matrix element form, this equation is

$$\frac{dA_{ij}}{dt} = A_{ik} \epsilon_{kjm} \omega_m, \tag{7}$$

where $\epsilon_{kjm}$ is the completely antisymmetric Levi-Civita symbol.

The effective axis and angle of rotation for time $t$ are extracted from $\mathbf{A}(t)$ by computing the trace of $\mathbf{A}$, tr $\mathbf{A}$, and constructing the intrinsic vector [9],

$$V_i = \epsilon_{ijk} A_{jk}. \tag{8}$$

These quantities are related to the effective angle of rotation $\alpha$, and a unit vector $\mathbf{n}$ along the effective axis of rotation by the formulas

$$\text{tr } \mathbf{A} = 1 + 2 \cos \alpha, \tag{9}$$

$$\mathbf{V} = 2\mathbf{n} \sin \alpha. \tag{10}$$

The handedness of the description of $\alpha$ with respect to $\mathbf{n}$ is fixed by eq (10) to agree with the common handedness of the coordinate systems $S$ and $S'$.

By differentiating eq (8) with respect to $t$, using eqs (7) and (10) to evaluate $dA_{jk}/dt$ and $d\mathbf{V}/dt$ respectively, and making use of the identity

$$\epsilon_{ijk} \epsilon_{msk} = \delta_{im} \delta_{js} - \delta_{is} \delta_{jm}, \tag{11}$$

we arrive at

$$2 \frac{d\mathbf{n}}{dt} \sin \alpha + 2\mathbf{n} \frac{d\alpha}{dt} \cos \alpha = \mathbf{B} \cdot \boldsymbol{\omega}, \tag{12}$$

where the dyadic $\mathbf{B}$ is given by

$$\mathbf{B} = (1 + 2 \cos \alpha)\mathbf{I} - \tilde{\mathbf{A}}. \tag{13}$$

Equation (12) is the starting point for computing the equations of motion for $\mathbf{n}$ and $\alpha$.

We first take the scalar product of eq (12) with $\mathbf{n}$. Since $\mathbf{n}$ is a unit vector, it is perpendicular

to its time derivative. Furthermore, from Euler's theorem [10]

$$\mathbf{n} \cdot \tilde{\mathbf{A}} \equiv \mathbf{A} \cdot \mathbf{n} = \mathbf{n}. \tag{14}$$

The result of dotting eq (12) with $\mathbf{n}$ is therefore

$$2\frac{d\alpha}{dt}\cos\alpha = (1+2\cos\alpha)\,\boldsymbol{\omega}\cdot\mathbf{n} - \boldsymbol{\omega}\cdot\mathbf{n},$$

$$= 2\boldsymbol{\omega}\cdot\mathbf{n}\cos\alpha.$$

Cancelling the common factor $2\cos\alpha$, we have

$$\frac{d\alpha}{dt} = \boldsymbol{\omega}\cdot\mathbf{n}. \tag{15}$$

Equation (15) states the compellingly simple, and intuitively almost obvious result that the instantaneous rate of change of the effective angle of rotation, is given by the component of the instantaneous angular velocity along the instantaneous effective axis of rotation.

We now insert the trigonometric representation of $\tilde{\mathbf{A}}$ [11],

$$\tilde{\mathbf{A}} = \mathbf{nn} + \mathbf{n}\times\mathbf{I}\sin\alpha - \mathbf{n}\times(\mathbf{n}\times\mathbf{I})\cos\alpha,$$

into eq (13) and use the result to evaluate $\mathbf{B}\cdot\boldsymbol{\omega}$ in eq (12). After replacing $d\alpha/dt$ *by* $\boldsymbol{\omega}\cdot\mathbf{n}$ and collecting similar terms we have

$$2\frac{d\mathbf{n}}{dt}\sin\alpha + \mathbf{n}\times\boldsymbol{\omega}\sin\alpha + \dot{\mathbf{n}}\times(\mathbf{n}\times\boldsymbol{\omega})(1+\cos\alpha) = 0. \tag{17}$$

This may also be written

$$\frac{d\mathbf{n}}{dt} = -\frac{1}{2}\left[\mathbf{n}\times\boldsymbol{\omega} + \mathbf{n}\times(\mathbf{n}\times\boldsymbol{\omega})\cot\frac{\alpha}{2}\right]. \tag{18}$$

There is no actual singularity at $\alpha = 0$ in eq (18) since for $\alpha = 0$, $\mathbf{n}$ and $\boldsymbol{\omega}$ are parallel.

Equation (15) and either of eqs (17) and (18) constitute the coupled equations of motion for the effective axis and angle in terms of the angular velocity. The inverse problem to that of obtaining these equations, is the problem of constructing the angular velocity from $\mathbf{n}$, $\alpha$, and their time derivatives. The construction is most readily carried out by formally solving eq (12) for $\boldsymbol{\omega}$. This requires the dyadic $\mathbf{B}^{-1}$ which is inverse to $\mathbf{B}$. To test whether the inverse exists, we compute det $\mathbf{B}$, which is most easily done by considering the special form of $\tilde{\mathbf{A}}(t)$ when $\mathbf{n}(t)$ instantaneously lies along one of the coordinate axes, say, the $\mathbf{b}_3$ axis. The result of the computation is

$$\det\mathbf{B} = 4\cos\alpha(\cos\alpha + 1). \tag{19}$$

Equation (19) shows that $\mathbf{B}^{-1}$ does not formally exist when $\alpha$ is $\pi/2$, $\pi$, or $3\pi/2$. In constructing $\mathbf{B}^{-1}$, we shall therefore assume that $\alpha$ is not equal to any of these special angles. Nevertheless, the final result for $\boldsymbol{\omega}$ will be seen to be well defined for all $\alpha$.

Using the trigonometric representation (16), and the identity

$$\mathbf{n}\times(\mathbf{n}\times\mathbf{I}) \equiv \mathbf{nn} - \mathbf{I},$$

we may write $\mathbf{B}$ in the form

$$\mathbf{B}=2[v^2\mathbf{I}-u^2\mathbf{nn}-uv\mathbf{n}\times\mathbf{I}], \tag{20}$$

where

$$u=\sin\frac{\alpha}{2}, \tag{21a}$$

$$v=\cos\frac{\alpha}{2}. \tag{21b}$$

We then assume an expansion of $\mathbf{B}^{-1}$ of the form

$$\mathbf{B}^{-1}=P\mathbf{I}+Q\mathbf{nn}+R\mathbf{n}\times\mathbf{I}, \tag{22}$$

and compute $P$, $Q$, and $R$ from the requirement

$$\mathbf{B}\cdot\mathbf{B}^{-1}=\mathbf{B}^{-1}\cdot\mathbf{B}=\mathbf{I}. \tag{23}$$

The results are

$$P=\frac{1}{2}, \tag{24a}$$

$$Q=\frac{u^2}{v^2-u^2}, \tag{24b}$$

$$R=\frac{1}{2}\frac{u}{v}. \tag{24c}$$

The expression for $\mathbf{B}^{-1}$ in terms of full angles is

$$\mathbf{B}^{-1}=\frac{1}{2}\left[\mathbf{I}+\frac{1-\cos\alpha}{\cos\alpha}\,\mathbf{nn}+\frac{\sin\alpha}{1+\cos\alpha}\,\mathbf{n}\times\mathbf{I}\right]. \tag{25}$$

If we now multiply eq (12) from the left by $\mathbf{B}^{-1}$, and note that $\mathbf{n}$ is perpendicular to its time derivative, and that $\mathbf{n}\times\mathbf{I}\cdot\mathbf{n}=0$, we arrive at

$$\boldsymbol{\omega}=\mathbf{n}\frac{d\alpha}{dt}+\frac{d\mathbf{n}}{dt}\sin\alpha+\mathbf{n}\times\frac{d\mathbf{n}}{dt}(1-\cos\alpha). \tag{26}$$

Equation (26) is the analog of eqs (1) in the parametrization of the rotation through $\mathbf{n}$ and $\alpha$, and gives the expansion of $\boldsymbol{\omega}(t)$ in the instantaneous mutually orthogonal trio of vectors $\mathbf{n}$, $d\mathbf{n}/dt$, and $\mathbf{n}\times d\mathbf{n}/dt$.

An interesting geometrical interpretation of the angle $\alpha$ emerges from eq (26). To arrive at this interpretation, we first construct the unit vectors which are parallel to $d\mathbf{n}/dt$ and $\mathbf{n}\times d\mathbf{n}/dt$ respectively. The lengths of these vectors are equal and, from eqs (15) and (18), are given by

$$\left|\frac{d\mathbf{n}}{dt}\right|=\left|\mathbf{n}\times\frac{d\mathbf{n}}{dt}\right|=\frac{1}{2}\sqrt{\omega^2-\left(\frac{d\alpha}{dt}\right)^2}\,\csc\frac{\alpha}{2}. \tag{27}$$

The quantity $[\omega^2 - (d\alpha/dt)^2]^{1/2}$ is the component of $\boldsymbol{\omega}$ perpendicular to $\mathbf{n}$. If $\beta$ is the angle between $\boldsymbol{\omega}$ and $\mathbf{n}$, we may write

$$\sqrt{\omega^2 - \left(\frac{d\alpha}{dt}\right)^2} = \omega \sin \beta. \tag{28}$$

We now define three mutually orthogonal unit vectors $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$ as follows:

$$\mathbf{q}_1 \equiv \frac{d\mathbf{n}/dt}{|d\mathbf{n}/dt|}, \tag{29a}$$

$$\mathbf{q}_2 \equiv \mathbf{n} \times \mathbf{q}_1, \tag{29b}$$

$$\mathbf{q}_3 \equiv \mathbf{n}. \tag{29c}$$

In terms of these vectors and the angle $\beta$ in eq (28), we may rewrite eq (26) in the form

$$\boldsymbol{\omega} = \omega(\mathbf{q}_1 \sin \beta \cos \frac{\alpha}{2} + \mathbf{q}_2 \sin \beta \sin \frac{\alpha}{2} + \mathbf{q}_3 \cos \beta). \tag{30}$$

From eq (30) we see that the angle $\alpha/2$ is the azimuthal angle of $\boldsymbol{\omega}$ with respect to the plane of $\mathbf{q}_1$ and $\mathbf{q}_3$, that is, the plane of $\mathbf{n}$ and $d\mathbf{n}/dt$. This interpretation is illustrated in figure 2, which shows the polar representation of $\boldsymbol{\omega}$ in the coordinate system formed by $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$.

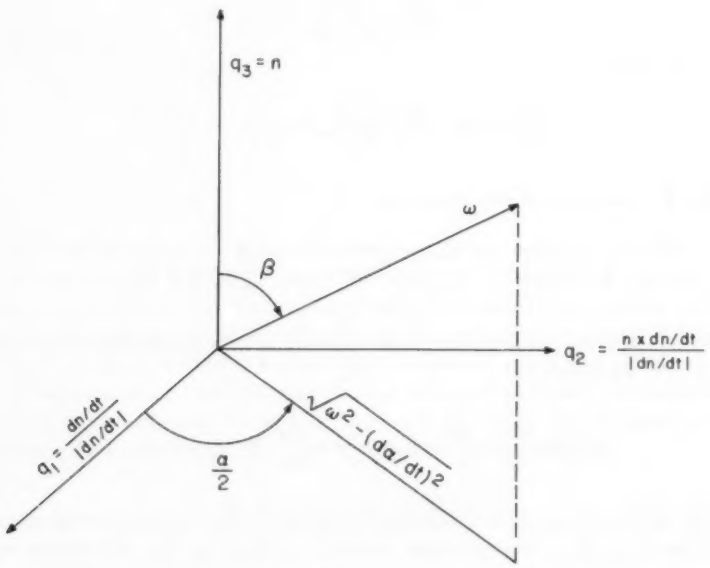


FIGURE 2. *Polar representation of the instantaneous angular velocity in a rotating coordinate system generated by a unit vector along the effective axis, and by the time derivative of that unit vector.*

The condition that the effective axis rotate rigidly with $S'$ is

$$\frac{d\mathbf{n}}{dt} = \boldsymbol{\omega} \times \mathbf{n}. \tag{31}$$

From eq (18), it is clear that in general, the motion of **n** departs significantly from a rigid rotation. The condition (31) is so restrictive that the only class of motions of $S'$ for which it is obeyed is that for which **n** is fixed in both $S$ and $S'$, with $S'$ consequently rotating about the fixed direction of **n**. This can be shown directly from eq (26), but is most easily shown from Euler's theorem. Differentiating eq (14), and making use of eq (6), we have

$$\mathbf{A} \cdot \left( \frac{d\mathbf{n}}{dt} - \boldsymbol{\omega} \times \mathbf{n} \right) = \frac{d\mathbf{n}}{dt}. \tag{32}$$

From eqs (31) and (32) we see that the condition that **n** rotate rigidly with $S'$ leads to

$$\frac{d\mathbf{n}}{dt} = 0. \tag{33}$$

Equation (26) then reduces to

$$\boldsymbol{\omega} = \mathbf{n} \frac{d\alpha}{dt}. \tag{34}$$

## 3. References


[1] Rose, M. E., Elementary Theory of Angular Momentum, Sec. 13 (John Wiley and Sons, Inc., New York, 1957).

[2] Murnaghan, Francis D., The Theory of Group Representations, Chap. 10, Sec. 1 (Johns Hopkins Press, Baltimore, 1938); (Dover Publications, Inc., New York, 1963).

[3] Moses, H. E., Irreducible Representations of the Rotation Group in Terms of Euler's Theorem. Nuovo Cimento **40A** (1965), 1120–1138.

[4] Moses, H. E., Irreducible Representations of the Rotation Group in Terms of the Axis and Angle of Rotation. Ann. Phys. (N.Y.) **37** (1966), 224–226.

[5] Moses, H. E., Irreducible Representations of the Rotation Group in Terms of the Axis and Angle of Rotation II. Orthogonality Relations Between Matrix Elements and Representations of Rotations in the Parameter Space. Ann. Phys. (N.Y.) **42** (1967), 343–346.

[6] Carmeli, M., Representations of the Three-Dimensional Rotation Group in Terms of Direction and Angle of Rotation. J. Math. Phys. **9** (1968), 1987–1992.

[7] Corben, H. C., and Stehle, P., Classical Mechanics, 2d ed., Sec. 49 (John Wiley and Sons, Inc., New York/London, 1960).

[8] Gelman, H., The Second Orthogonality Conditions in the Theory of Proper and Improper Rotations, I. Derivation of the Conditions and of Their Main Consequences. J. Res. Nat. Bur. Stand. (U.S.), 72B (Math. Sci.) No. 3 (July–Sept, 1968), 229–237.

[9] Gelman, H., The Second Orthogonality Conditions in the Theory of Proper and Improper Rotations, II. The Intrinsic Vector. J. Res. Nat. Bur. Stand. (U.S.), 73B (Math. Sci.), No. 2 (April–June, 1969), 125–138.

[10] Goldstein, H., Classical Mechanics, Sec. 4–6 (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1950).

[11] Page, L., Introduction to Theoretical Physics, 3d ed., Sec. 18 (D. Van Nostrand Company, Inc., New York, 1952).


(Paper 75B3&4–356)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# A Numerical Solution of Integral Equation of the First Kind Applied to Slit Correction in Small-Angle X-Ray Scattering

**Jacob Mazur**

**Institute for Materials Research, National Bureau of Standards, Washington, D.C. 20234**

**(November 2, 1971)**

A novel approach to the problem of the slit correction in small-angle x-ray scattering is presented, based on a matrix inversion method. The integral equation for the slit correction is written as a Volterra equation of the first kind. This equation is reduced to a system of simultaneous equations, expressed in matrix form. The order of the matrix is equal to the number of experimentally determined points. To obtain these equations, one has to expand the unknown functions in Taylor series around each point to be subsequently determined. There is, however, a difficulty inherent in this method due to the fact that most of the series expansions of the function to be determined lead to a strong numerical instability. However, a general method is developed, which enables us to find shifting operators leading to numerically stable systems of equations. The "unsmearing" of the experimental data is then performed by standard matrix inversion procedures.

Key words: Conditions for convergence of solutions; integral equation of first kind; shifting operators; slit correction; slit weighting function; small-angle x-ray scattering; strong numerical instability.

## 1. Introduction

In a previous paper [1][1] methods for solving the integral equation for slit correction in small angle x-ray scattering experiments were proposed. This equation relates the experimentally determined intensity of scattered rays to the true scattering intensity as follows:

$$\tilde{I}(x) = 2\int_x^\infty \frac{yI(y)W(\sqrt{y^2-x^2})dy}{\sqrt{y^2-x^2}}. \tag{1}$$

In eq (1), $\tilde{I}(x)$ is the experimentally determined intensity of the scattered light, $I(x)$ is the true scattering intensity, $x$ is proportional to the scattering angle [1], and $W(x)$ is the experimentally determined slit weighting function. The dependence of $W$ on the slit collimation system is widely discussed in references [2–5]. It may be given either in analytic form or only for discrete values of its argument.

$I(x)$ is to be calculated from eq (1), which is a Volterra integral equation of the first kind. In Reference [1] explicit and implicit solutions have been proposed. The explicit solution is based on transforming eq (1) into a Volterra equation of the second kind, which has for its inhomogeneous part a constant term, and whose kernel involves an integrated expression of the derivative of the experimentally determined slit weighting function $W$. This solution requires numerical differentiation of the experimental data, a process which may introduce serious errors. Differentiation is not required in the implicit solution method. Rather, the experimental data enter into the solution directly.

*AMS Subject Classification:* Primary 4510.

[1] Figures in brackets indicate the literature references at the end of this paper.

This paper will deal only with the implicit solution method. In that case, eq (1) is first reduced to a system of simultaneous linear equations. These equations are then solved by inversion of the resulting triangular matrix. The solution leads to calculated approximate values of $I(x)$, given on a set of $N$ discrete points $x=0$, $x=\Delta x$, . . ., $x=(N-1)\Delta x$. These are the same points for which values of $\tilde{\tilde{I}}(x)$ must be given by experiment.

One of the difficulties in applying this procedure directly is that it can lead to numerical instability. However, one can find an interpolating polynomial for $I(x)$ such that a system of simultaneous equations is obtained which is numerically stable. In this paper will explore ways to find an accurate interpolating polynomial for the purpose of our computation. However, the problem of finding a suitable polynomial can be altogether avoided if another, indirect method is employed. That method consists in first obtaining a zeroth-order approximation to the solution, based on a single-step method and which is known to be always numerically stable. Subsequent improvements to the solution are then computed, using standard perturbation methods to obtain subsequent corrections to the zeroth-order approximation.

The first of these two approaches involves application of multi-step methods which are based on existence of polynomial which interpolates the function $I(x)$ at discrete points on a given interval of $x$. The second approach is based on a single-step method, which serves as a starting point for the perturbation method. Both approaches are included in a computer program that solves the problem of slit correction in small angle x-ray scattering. Four different choices for the weighting function ($W$ in eq (1)), have been used in that program:

1. Constant weighting function (infinite slit)
2. Gaussian
3. Trapezoid, and
4. Data weighting function.

The details of these computations and an analysis of the results will appear in an NBS Technical Note, to be published. In this Note, the Fortran program that performs the required calculations will be described in detail.

## 2. Algebraic Formulation of the Problem

The kernel of the integral eq (1) has a singularity at $y=x$. In order to deal with this singularity, we follow the procedure employed in solving Abel's integral equation. Both sides of eq (1) are multiplied by $x\,dx/\sqrt{x^2-u^2}$ and integrated over $x$ from $u$ to infinity. The transformed equation, after exchanging the orders of integration is

$$\int_u^\infty \tilde{\tilde{I}}(x)G(u,x)\,dx=\int_u^\infty I(x)K(u,x)\,dx \tag{2}$$

with

$$K(u,x)=2x\int_0^{\sqrt{x^2-u^2}}\frac{W(t)\,dt}{\sqrt{x^2-u^2-t^2}} \tag{2a}$$

and

$$G(u,x)=x/\sqrt{x^2-u^2}$$

The total number of experimental data for $\tilde{\tilde{I}}(x)$ is $N$, which are given at equally spaced intervals $\Delta x$. The range of $x$ is determined in such a way that putting $\tilde{\tilde{I}}(x)$ equal to zero for all $x \geq N\Delta x$ does not affect the solution. In the physical problem we are investigating, the assumption that $\tilde{\tilde{I}}(x)=0$ for $x \geq N\Delta x$ implies that $I(x)=0$ for the same values of $x$. Therefore, both the upper limits of integrals in (2) are replaced by $N\Delta x$. Thus, the function $\tilde{\tilde{I}}(x)$ represents a discrete set of function values $\tilde{\tilde{I}}_i$, $i=1,2,\ldots,N$. Both $\tilde{\tilde{I}}(x)$ and $I(x)$ may therefore be viewed as $N$-dimensional vectors. The first components $\tilde{\tilde{I}}_1$ and $I_1$ of these vectors correspond to $u=0$; $i=N$ corresponds to $u=(N-1)\Delta x$.

It should be remarked, that consistent with the above restrictions on the functions $\tilde{\tilde{I}}(x)$ and $I(x)$, the method of solution presented in this paper is not necessarily restricted to the particular representations of $K(u, x)$ and $G(u, x)$, which in the problem being investigated are given by eqs (2a) and (2b). Hopefully, the method presented here is general and might be applied to other physical problems which can be simulated by eq (2).

The left-hand side of eq (2) is calculated from known quantities. The results are represented by the vector **h** with components

$$h_i = \int_{(i-1)\Delta x}^{N\Delta x} \tilde{\tilde{I}}(x) G(u, x)\, dx. \tag{3}$$

In our particular problem, $G(u, x)$ is singular at $x = u$ and is given by eq (2b). For this reason, a direct numerical evaluation of the integral (3) will not be practical, since we will have to subdivide the integrand very closely at its lower end in order to achieve sufficient accuracy. To evaluate (3) we replace the integral by summation as follows:

$$\int_u^{\infty} \frac{x\tilde{\tilde{I}}(x)\,dx}{\sqrt{x^2-u^2}} \cong \sum_{j=i}^{N} \int_{(j-1)\Delta x}^{j\Delta x} \frac{xI(x)\,dx}{\sqrt{x^2-(i-1)^2\Delta x^2}}. \tag{4}$$

Next, $\tilde{\tilde{I}}(x)$ is expanded in a Taylor series around $x = j\Delta x$:

$$h_i = \sum_{j=i}^{N} \int_{(j-1)\Delta x}^{j\Delta x} \left[ I(j\Delta x) + (x - j\Delta x)A_1 + \frac{1}{2}(x-j\Delta x)^2 A_2 + \ldots \right] \frac{x\,dx}{\sqrt{x^2-(i-1)^2\Delta x^2}} \tag{5}$$

$A_1$, $A_2$, etc., are the first, second, etc. derivatives of $\tilde{\tilde{I}}(x)$ evaluated at $x = j\Delta x$. These derivatives can be approximated by one of several difference relations. We may use the following finite difference expression:

$$A_1 = (\tilde{\tilde{I}}_{j+1} - \tilde{\tilde{I}}_j)/\Delta x, \qquad A_2 = (\tilde{\tilde{I}}_{j+2} - 2\tilde{\tilde{I}}_{j-1} + \tilde{\tilde{I}}_j)/(\Delta x)^2.$$

Only two derivatives are used in the Taylor series in (5); $\tilde{\tilde{I}}_{N+1}$ and $\tilde{\tilde{I}}_{N+2}$ are set equal to zero. After performing the indicated integrations and rearranging eq (5), the result is written in the form

$$h_i = \sum_{j=i}^{N} \gamma_{i,j} \tilde{\tilde{I}}_j; \qquad \tilde{\tilde{I}}_j = 0 \text{ for } j > N. \tag{6}$$

or $\mathbf{h} = \mathbf{A}\tilde{\tilde{\mathbf{I}}}$. **A** is an upper triangular matrix with entries the coefficients $\gamma_{i,j}$. The matrix elements $\gamma_{i,j}$, have different representations depending on the following three cases: (1) $j - i = 0$, (2) $j - i$ even, and (3) $j - i$ odd. In addition, in order to close the system, one has to consider, separately, the cases $j = N$ and $j = N - 1$. This is because the last integral, if taken from $x = (N-1)\Delta x$, cannot include second-order differences if forward extrapolation is employed. Also for the case $i = 1 (u = 0)$ we have to avoid the expression $(i-1)\log(i-1)$. Details will be given in an NBS Technical Note describing the computer program.

We now expand the r.h.s. of the integral equation (2) in terms of the discrete values $I_i$, as

$$\int_{(i-1)\Delta x}^{N} I(x)K(u,x)\,dx = \sum_{j=i}^{N} b_{i,j} I_j; \quad I_j = 0 \quad \text{for} \quad j > N. \tag{7}$$

to obtain a set of simultaneous equations

$$\sum_{j=1}^{N} b_{i,j} I_j = h_i \tag{8}$$

$b_{i,j}(i=1, 2, \ldots, N;\ j=i, i+1, \ldots, N)$ are coefficients which we have not yet determined. These coefficients depend on the functional form of the kernel $K(u, x)$, and, therefore, on the assumed form for the weighting function $W$ of eq (1). In addition, they also depend on the coefficients of the interpolating polynomial which approximates the function $I(x)$.

The solution of eq (8) can be expressed in matrix form:

$$\mathbf{I}=\mathbf{B}^{-1}\mathbf{h}. \tag{9}$$

$\mathbf{I}$ is the $N$-dimensional solution vector. $\mathbf{h}$ is an $N$-dimensional vector whose elements were already computed from the experimental data and the indicated numerical integrations. The inversion of an upper triangular matrix $\mathbf{B}$ can be rapidly performed by a back substitution [5], starting with the last equation. With this method, the elements $c_{i,j}$ of the inverse matrix $\mathbf{B}^{-1}$ are obtained by first calculating the elements $d_{i,j}$ from the following recurrence equations:

$$d_{i,i}=1,$$

$$d_{i,j}=\sum_{k=i}^{k=j-1}-\left(\frac{b_{k,j}}{b_{i,k}}\right)d_{i,k} \tag{10}$$

and then

$$c_{i,j}=d_{i,j}/b_{i,j}.$$

For certain forms of $K(u, x)$, the kernel of the r.h.s. of eq (2), there is a repetitious pattern in the representations of the elements $d_{i,j}$. In this case, the number of different elements in the matrix $\mathbf{B}$ is less than $(1/2)N(N+1)$. In particular, if $K(u, x)$ is independent of $u$ (e.g., when $W$ of eq (2a) is constant and therefore, $K(u, x)=2x$) the elements of the matrix $\mathbf{B}$ (after multiplication by the diagonal matrix

$$\left\{\frac{1}{i}\delta_{ij}\right\}$$

depend only on $|j-i|$. This has an important practical aspect: computer computations with large matrices require storage capacity which is often not available. It is then necessary to employ vectors $\mathbf{b}_j$ (formed separately for each row of the matrix $\mathbf{B}$), or, to compute $b_{i,j}$ each time it appears in computations, separately. The problem is, that these computations will be repeated many times. However, for this case, only the first row of the matrix, in form of a vector $\mathbf{b}_N$ needs to be computed, and the problem of storing large matrices is avoided.

Any algorithm in which the calculated value of the increment $I_i-I_{i+1}$ depends only on $I_{i+1}$ and on the step $\Delta x$, is called a one-step method. In other words, an expression for $I_i$ may be written in a closed form involving explicitly only $I_{i+1}$. Of course, $I_{i+1}$ depends, in its turn, on $I_{i+2}$, etc. In a multi-step method of multiplicity $l$, the increment $I_i-I_{i+1}$ depends also on $I_{i+2}, \ldots, I_{i+l}$. It therefore follows, that whenever one approximates $I(x)$ in eq (1) by an interpolating polynomial of order higher than one, the method of calculating $I_i$ in terms of already computed $I_{i+1}$, $I_{i+2}$, etc., will be defined as a multi-step method of multiplicity which is equal to the order of the interpolating polynomial. In a multi-step method, a special starting procedure is required to find the "initial" or "starting" values of $I_N, I_{N-1}, \ldots, I_{N-l}$. We calculate these starting values by trapezoid rule as follows:

$$h_{N-1}=\frac{1}{2}\int_{(N-2)\Delta x}^{(N-1)\Delta x}(I_N+I_{N-1})K[(N-2)\Delta x, x]dx+\frac{1}{2}I_N\int_{(N-1)\Delta x}^{N\Delta x}K[(N-2)\Delta x, x]dx$$

$$2h_{N-l}=I_{N-l}\int_{(N-l-1)\Delta x}^{(N-l)\Delta x}K[(N-l-1)\Delta x, x]dx+I_{N-l+1}\int_{(N-l-1)\Delta x}^{(N-l+1)\Delta x}K[(N-l+1)\Delta x, x]dx$$

$$\ldots+I_N\int_{(N-2)\Delta x}^{N\Delta x}K[(N-l-1)\Delta x, x]dx$$

We have $l$ equations and $l+1$ unknowns. However, a physical assumption was made that setting $h=0$ for $i>N$, implies that $I_i=0$ for $i>N$. For this reason,

$$I_N=2h_N \Big/ \int_{(N-1)\Delta x}^{N\Delta x} K[(N-1)\Delta x, x]dx.$$

Let us divide the integral $h_i$

$$h_i=\int_{(i-1)\Delta x}^{(N-1)\Delta x} I(x)K[(i-1)\Delta x, x]dx$$

into smaller integrals, each with a range of integration equal to $l\Delta x$ (except for the last integral). For this purpose, let us define

$$\kappa_{i,j}=\int_{(j-1)\Delta x}^{N\Delta x} I(x)K[(i-1)\Delta x, x]dx \qquad (j=i, i+1, \ldots, N-1) \tag{11}$$

so that $\kappa_{i,i}=h_i$.
$h_i$ is the sum of differences of the $\kappa_{i,j}$ as follows:

$$h_i=(\kappa_{i,i}-\kappa_{i,i+l})+(\kappa_{i,i+l}-\kappa_{i,i+2l})+\ldots$$

(In this expression, the number of bracketed terms is equal to the nearest integer not exceeding $(N-i)/l$. In addition, a term is added to close the system so that $h_i=\kappa_{i,i}$.) Next, in each of the difference equations

$$\kappa_{i,j}-\kappa_{i,j+l}=\int_{(j-1)\Delta x}^{(j+l-1)\Delta x} I(x)K[(i-1)\Delta x, x]dx \tag{12}$$

we replace $I(x)$ by an interpolating polynomial having the values $I_k=I(x_k)$ on a set of points $x_k$. These points might be taken as $x_k=(j-1)\Delta x, \ldots, (j-l-1)\Delta x$. In forward interpolation, the interpolating points are the $p$ points $j-1+m, j+m, \ldots, j+m+p-1$, with $p \geq l$. $m$ denotes the point from which the differences are computed. We will always have $m<l$, so that the point of reference lies within the range of the integral (12). The polynomial $Q(x)$ which approximates and replaces $I(x)$ in eq (12), in terms of forward differences [6], is

$$Q(x)=\sum_{q=0}^{p}\binom{(x/\Delta x)-j-m+1}{q}\Delta^q I_{j+m}. \tag{13}$$

In eq (13), $\Delta$ is a forward difference operator. Thus, for $q=1$, $\Delta I_{j+m+1}=I_{+m+1}-I_{j+m}$, etc. $\Delta^r I_{j+m}=\Delta I_{j+m}$.

Substituting eq (13) in eq (12), integrating and rearranging, the result is formally written as

$$\kappa_{i,j}-\kappa_{i,j+l}=\alpha_{i,j}I_{j+m}+\alpha_{i,j+1}I_{j+m+1}+\ldots+\alpha_{i,j+p}I_{j+m+p} \tag{14}$$

$\alpha_{i,j+k}(k=0, 1, \ldots, p)$ are the coefficients resulting from the above steps taken in order to approximate eq (12) by eq (14).

In the formal solution, written in matrix form as $\mathbf{I}=\mathbf{B}^{-1}\mathbf{h}$, the elements of the matrix $\mathbf{B}$ can be expressed in terms of $\alpha_{i,j}$ of expansion (14).

Before proceeding, let us stop and see what will happen to our computations with $I(x)$ of eq (12) being replaced by polynomial $Q(x)$, for the specific case of $m=0$ and $l=p=2$. In this case,



449-764 OL - 72 - 6

all three interpolated values of $I_i$ are within the range of the integral of eq (12). Let us assume tha $K(u, x) \propto x$. This is the case when the weighting function $W$ is constant ($W$ = constant is descrip tive of an infinitely-wide slit in small angle x-ray scattering). For $i \gg 1$, the elements of a typica row of the **B** matrix are found to be

$$b_{i,i}=\frac{i}{3};\ b_{i,i+1}=\frac{4}{3}\ (i+1);\ b_{i,i+2}=\frac{2}{3}\ (i+2).$$

The coefficients $\alpha_{i,j}$ of eq (13) are:

$$\alpha_{i,i}=\frac{i}{3};\ \alpha_{i,i+1}=\frac{4}{3}\ (i+1);\ \alpha=\frac{1}{3}\ (i+2).$$

The absolute values of the elements of the inverse of the **B** matrix can be shown to grow exponentially as one moves away from the diagonal elements. Consequently, the numerical calculations are affected by strong numerical instability.

## 3. Stability Conditions for Integral Equations of First Kind

### 3.1. General Theory

Consider first the difference eq (14) written for $\kappa_{i,i}-\kappa_{i,i+l}$ and for $p=l$. The associated homogeneous equation, obtained by setting

$\kappa_{i,i}-\kappa_{i,i+l}=0$, is

$$\alpha_{i,i}I_{i+m}+\alpha_{i,i+1}I_{i+m+1}+\ .\ .\ .\ +\alpha_{i,l}I_{i+m+l}=0. \tag{15}$$

Solution of (15) can be found for every choice of starting values $I_N, I_{N-1}, \ldots, I_{N-l+1}$. (These are computed using the trapezoid rule and the one-step method.) In particular, if these starting values are all equal to zero, then there corresponds a trivial solution $I_i=0$ for every $i$.

Consider first a simplified homogeneous equation for the starting values of **I**, with constant coefficients, replacing eq (15),

$$\alpha_0 I_{N-l}+\alpha_1 I_{N-l+1}+\ .\ .\ .\ +\alpha_l I_N=0. \tag{16}$$

We determine the conditions for convergence of solution of eq (16) for $N\to\infty$, while retaining finite value for $N\Delta x$. That is, the intervals between successive points becomes smaller and smaller as $N$ increases, and, as $N\to\infty$, $\Delta x\to 0$. If the multi-step method for solving eq (16) converges, then it will converge also for the special value of $I_N=0$.
Therefore,

$$\lim_{\Delta x\to 0} I_\mu=0, \qquad \mu=N,\quad N-1,\ .\ .\ .,\quad N-l \tag{17}$$

is a condition for convergence of multi-step method for solving eq (16). Accordingly, if condition (17) is not satisfied, the solution of the general homogeneous eq (15) will diverge in the limit of $\Delta x\to 0$. This divergence implies the divergence of the solution of the inhomogeneous eq (14). The procedure to determine the conditions for convergence of solution of eq (15) in the limit of $\Delta x\to\infty 0$ (or $N\to\infty$), is identical with the one employed in connection with numerical solution of difference-differential equations [7], and proceeds as follows: The homogeneous eq (15), with constant coefficients, is expressed as:

$$\alpha_i I_{i+m}+\alpha_{i+1}I_{i+m+1}+\ .\ .\ .\ +\alpha_{i+l}I_{i+m+l}=0. \tag{18}$$

In eq (18), we substitute

$$I_{i+m+l}=\eta^{-(i+m+l)}.$$

Equation (18) becomes after division by common factor $\eta^{-(i+m+l)}$:

$$\alpha_i\eta^l+\alpha_{i+1}\eta^{l-1}+\ .\ .\ .\ +\alpha_{i+l}=0. \tag{19}$$

Therefore, $\eta$ must be a root of the characteristic polynomial $\emptyset(\eta)$:

$$\emptyset(\eta)=\alpha_i\eta^l+\alpha_{i+1}\eta^{l-1}+\ .\ .\ .\ +\alpha_{i+l}=0. \tag{20}$$

We require that the multi-step method be convergent. The necessary condition for the convergence of the general multi-step method for our integral equation of the first kind is the same as the condition for the convergence of difference equations of *first* order [8], and is stated by the following:

THEOREM: A necessary condition for the convergence of the multi-step solution (14) for integral equation (1) is that the moduli of all roots of the associated polynomial $\emptyset(\eta)$ do not exceed unity, and that the root of modulus 1 is simple.

The condition imposed on $\emptyset(\eta)$ is also the condition of stability [9]. Violation of this condition leads to "strong" numerical instability. This instability manifests itself by the error growth at a fixed point $x=(i-1)\Delta x$ as the spacing between the points decreases. Strong instability is therefore characterized as a pointwise phenomenon, as it involves error growth at fixed points as the interval between them decreases. In contradistinction to this, a "weak" instability, to be discussed subsequently in connection with even-odd oscillations in computed values of $I_i$, is associated with the truncation error, and exists even when the method employed is convergent. Since the weak instability is associated with the total range of $x$, while keeping the interval between successive points constant, this instability is also known as a stepwise phenomenon [7, 10].

Let us demonstrate the concept of strong instability by a concrete example. Assume that $l=2$ and $m=0$. Also, let $K(u,x)\propto x$, so that $\kappa_{i,j}-\kappa_{i,j+2}=h_i-h_{i+2}$. We substitute eq (13) into eq (12) and integrate. Since the r.h.s. of eq (12) is a function of $j$ only, $\alpha_{i,j}$, the elements of the matrix $\mathbf{B}$, are shown to be equal to $j\bar{\alpha}_{|j-i|}$ for $j\geq i$ and to zero for $j<i$. $\bar{\alpha}_{|j-i|}$ is a function of $j-i$ only. If, in addition, $p=l$, then, for $i\gg 1$,

$$h_i-h_{i+2}=\alpha_0 I_i+\alpha_1 I_{i+1}+\alpha_2 I_{i+2} \tag{21}$$

where

$$\alpha_0=\alpha_{i,i}\cong i\bar{\alpha}_0,\ \alpha_1=\alpha_{i,i+1}\cong i\bar{\alpha}_1$$

and

$$\alpha_2=\alpha_{i,i+2}\cong i\bar{\alpha}_2.$$

For a particular case when $\alpha_0=\alpha_2$ (as in the example discussed at the end of chap. 2), the elements $b^*_{i,j}$ of the inverse matrix $\mathbf{B}^{-1}$ can be expressed directly in terms of the roots of the characteristic polynomial $\emptyset(\eta)$, eq (19). One obtains

$$ib^*_{i,j}=\left[\frac{1}{2\alpha_0}\left(-\alpha_1+\sqrt{\alpha_1^2-4\alpha_1\alpha_0}\right)\right]^k+\left[\frac{1}{2\alpha_0}\left(-\alpha_1-\sqrt{\alpha_1^2-4\alpha_1\alpha_0}\right)\right]^k \tag{22}$$
$$(j>i)$$

and $b^*_{i,i}=1/i$.

In eq (22), $k=j-i$. When $\alpha_2 \neq \alpha_0$, the relationships between $b^*_{i,j}$ and the roots of eq (19) are more involved. However, for sufficiently large $k$, the leading term in the expression for $b^*_{i,k+i}$ can be shown to be equal to

$$\left|\frac{-\alpha_1}{2\alpha_0}\right|^k$$

For the case of $K(u, x) \propto x$, the following equation is derived

$$I_i - I_{i+2} = \sum_{j=i}^{N} b^*_{i,j}[h_j - (i/(i+2))h_{j+2}]. \tag{23}$$

When $k$ becomes infinitely large $I_i$ will remain finite only if $\lim_{(j-i)\to\infty} b^*_{i,j}$ converges. This happens only when both expressions in parentheses in eq (22) do not exceed 1 in their absolute values. This implies that the moduli of the roots of the equation

$$\phi(\eta) = \eta^2 + \alpha_1\eta + \alpha_2 = 0$$

do not exceed unity. This simple example demonstrates the intimate connection between the root condition and the stability of the multi-step method.

In the general theory of stability of the difference differential equations, the necessary and sufficient conditions for the convergence of the multi-step method require, in addition to the root condition, also the satisfaction of the consistency condition [9]. This condition requires that, in order that an expansion in discrete variables (e.g., our expansion eq (14)) be a good approximation to the original equation which it simulates, certain restrictions must be imposed on the coefficients in such expansion. In our problem the consistency condition requires that the sum of the coefficients in eq (18) be equal to $l$:

$$\sum_{r=1}^{r=i+l} \alpha_r = l. \tag{23}$$

Since this condition is always satisfied, the root condition becomes both the necessary and the sufficient condition for the convergence of the multi-step method. We also see why the method for the determination of the starting values of $I_i$, based on a one-step method, is always stable and convergent: The polynomial $\phi(\eta)$ is then of the first order with its root equal to 1.

### 3.2. Determination of Stability Conditions

Our purpose is to obtain an approximation to the integral

$$\int_{(i-1)\Delta x}^{(i+l-1)\Delta x} I(x)K[(i-1)\Delta x, x]dx = \kappa_{i,i} - \kappa_{i,i+l} \tag{24}$$

in terms of finite differences, which satisfies the stability conditions. These differences can be referred to the point $x=(i-1)\Delta x$, to the point $x=i\Delta x$, and so on. Let the point of reference be located at $x=(i+m-1)\Delta x$, $m=0, \ldots, l$. We define a shift (displacement) operator $\mathbf{E}$ by

$$\mathbf{E}I_j = I_{j-1}. \tag{25}$$

Applying the shifting operator to the point of reference, one obtains

$$I(x) = \mathbf{E}^{-t}(I_{i+m}). \tag{26}$$

$t$ is a new variable, defined by $t=x/\Delta x-(i+m-1)$, and it measures the distance from the point of reference. Therefore,

$$\kappa_{i,j}-\kappa_{i,i+l}=\Delta x\int_{-m}^{(l-m)}K(i^*,i^*+t^*+m^*)\,\mathbf{E}^{-t}[I_{i+m-1}]dt. \tag{27}$$

In eq (27) and in subsequent equations, $i^*=(i-1)\Delta x$, $t^*=t\Delta x$, and $m^*=m\Delta x$ and $K(i^*, i^*+t^*+m^*)=K[(i-1)\Delta x, (i+t+m-1)\Delta x]$. The operator $\mathbf{E}$ can be expressed in terms of one of the three difference operators, $\boldsymbol{\Delta}$, $\boldsymbol{\nabla}$, and $\boldsymbol{\delta}$ which are defined by

$\boldsymbol{\Delta} I_j=I_{j+1}-I_j$ (forward difference operator),

$\boldsymbol{\Delta} I_j=I_j-I_{j-1}$ (backward difference operator),

$\boldsymbol{\delta} I_j=I_{j+1/2}-I_{j-1/2}$ (central difference operator).

These operators are related to $\mathbf{E}$ through

$$\boldsymbol{\Delta}=1-\mathbf{E},\ \boldsymbol{\Delta}=\mathbf{E}^{-1}-1,\ \boldsymbol{\delta}=\mathbf{E}^{1/2}-\mathbf{E}^{-1/2}.$$

For the backward difference operator $\boldsymbol{\Delta}$ we have the following eq:

$$\kappa_{i,i}-\kappa_{i,i+l}=\Delta x\int_{-m}^{l-m}K(i^*,i^*+t^*+m^*)[(1-\boldsymbol{\Delta})^{-t}I_{i+m-1}]dt. \tag{28}$$

For forward difference operator $\boldsymbol{\Delta}$ we have

$$\kappa_{i,i}-\kappa_{i,i+l}=\boldsymbol{\Delta x}\int_{-m}^{l-m}K(i^*,i^*+t^*+m^*)\ [(1+\boldsymbol{\Delta})^{t}I_{i+m-1}]dt. \tag{29}$$

In order to derive an equation for $\kappa_{i,i}-\kappa_{i,i+l}$ for central differences in terms of difference operator $\boldsymbol{\delta}$, we have for express $\mathbf{E}$ in terms of $\mathbf{d}$. Let $\mathbf{E}=e^{\mathbf{D}}$, and, therefore, $\boldsymbol{\delta}=2\sinh(\mathbf{D}/2)$. Thus,

$$\mathbf{D}=2\sinh^{-1}(\boldsymbol{\delta}/2)=\boldsymbol{\delta}-\frac{\boldsymbol{\delta}^3}{2^2 3!}+\frac{3^2\boldsymbol{\delta}^5}{2^4 5!}+\ldots\cdot$$

$$\mathbf{E}^{-t}=1-t\left(\boldsymbol{\delta}-\frac{\boldsymbol{\delta}^3}{3!2^2}+\ldots\right)+\frac{t^2}{2}\left(\boldsymbol{\delta}-\frac{\boldsymbol{\delta}^3}{3!2^2}+\ldots\right)^2\cdot \tag{30}$$

Retaining terms involving no higher operators than $\boldsymbol{\delta}^2$, we have $\mathbf{E}^{-t}=1-\boldsymbol{\delta}t+t^2\boldsymbol{\delta}^2/2$.
Therefore,

$$\kappa_{i,i}-\kappa_{i,i+2}=\Delta x\int_{-m}^{l-m}K(i^*,i^*+t^*+m^*)\left[\left(1-\boldsymbol{\delta}t+\frac{t^2\boldsymbol{\delta}^2}{2}\right)I_{i+m-1}\right]dt\cdot \tag{31}$$

Let us now investigate a few cases with simplified kernels, and find out which ones lead to numerically stable solutions. For the kernel $K(i^*, i^*+t^*+m^*)$ we again adopt the form which is descriptive of the constant weighting function in small-angle x-ray scattering, that is, $K(i^*, i^*+t^*+m^*)=(i+t+m-1)\Delta x$. For the first case let us take the form is which the approximating polynomial agrees with $I(x)$ *at* $x=(i-1)\Delta x, \ldots, (i+l-1)\Delta x$. We take $l=2$, and $m=2$. In order that all points will fall within the range of the integral, we must use the $\nabla$ (backward) operator. Consider $i\gg 1$. The root-determining equation is $1/3(x^2+4x+1)=0$. (The coefficient of $x^2$ is the coefficient of $I_i$, the coefficient of $x$ is the coefficient of $I_{i+1}$, and the last term is the coefficient of $I_{i+2}$.) We detect immediately that this form leads to a divergent solution. (Root condition is not satisfied.)



449-764 OL - 72 - 7

Let us now consider the case with central differences centered about the mid-point of the integrand. Let $l=2$, and $m=1$. From eq (31), the following root-determining equation is derived:

$$1/3x^2+4/3x+1/3=0 \tag{32}$$

and the method is unstable. It is interesting to notice that this particular case, which is the most obvious one since it is symmetric and does not involve odd powers of $t$ (in the integrand of eq (31), leads to an equation which is equivalent to the extended Simpson rule for numerical integration. The same result would be obtained if we evaluated the integral as

$$\int_{i\Delta x}^{N\Delta x} f(i\Delta x,x)dx=1/3(f_{i,i}+4f_{i,i+1}+2f_{i,i+2}+\ .\ .\ .)\Delta x \tag{33}$$

with $f(i\Delta x,x)=K(i\Delta x,x)I(x)$, etc.

At this point, we are led to the conjecture that the most straightforward cases, namely the ones that involve points located within the range of the integral, lead to unstable solutions. If the above restriction is removed, stability conditions may be satisfied. For example, in the case with $m=2$, (the finite differences refer to the point $x=(i+1)\Delta x$) we have, for $l=2$ and the $\Delta$ operator, two fixed points which we located outside the range of the integral. The root condition can be shown to be satisfied.

The method that we find to be the best for our purposes is the one in which the $\Delta$ operator is used and the absolute values of the coefficients of $\Delta^r, r=1,2,\ .\ .\ .\ ,p$, are made as small as possible. This can be achieved by selecting the point of reference to be located in the middle of the integrand, i.e., by taking $m=l/2$. In this case, one is justified in retaining only the first few terms in the expansion of (27). We have, for $l=2$,

$$\kappa_{i,i}-\kappa_{i,i+2}=\Delta x\int_{-1}^{1}K(i^*,i^*+t^*)\left[1-t\Delta+\frac{t}{2}\ (t-1)\Delta^2-\ .\ .\ .\right]I_i dt. \tag{34}$$

For $K(i^*,i^*+t^*)=(i+t)\Delta x$, and $i\gg 1$, the coefficients of $I_{i+j}(j=0,1$ or $2)$ in the expansion (34), after integration, are identical to the ones employed in Nÿstrom's method [11]. Since the root determining equation is

$$7/3x^2-2/3x+1/3=0, \tag{35}$$

the root conditions are satisfied. We found that, of all the multi-step methods, the above method is the most accurate for our purposes. Part of the computer program for the slit correction in small-angle x-ray scattering is based on this algorithm.

In conclusion, to insure stability, the differences in the expansion of the shifting operator should be referred to the point $(I_i)$ to be calculated and forward differences should be used. We also conjecture that the central difference method, which is the most accurate one in solving ordinary difference differential equation, is always numerically unstable.

The following qualitative considerations indicate that the stability conditions are only slightly altered by the actual form of $K[(i-i)\Delta x, x]$ and that a non-constant weighting function relaxes the condition for the attainment of the stability conditions. Consider $K(u, x)=xF(\sqrt{x^2-u^2})$, (which is the form used in actual computations of slit corrections). $F$ is related to the weighting function $W$ by

$$F(\sqrt{x^2-u^2})=\int_0^{\sqrt{x^2-u^2}}\frac{W(y)dy}{\sqrt{x^2-u^2-y^2}} \tag{36}$$

Now, let us form a matrix of $b_{i,j}$, where $b_{i,j}$ are the coefficients in the approximation of

$$\int_u^{N\Delta x} xI(x)F(\sqrt{x^2-u^2})dx \text{ by an expansion } \sum_{j=i}^{N} b_{i,j}I_j.$$

Consider, for example, a trapezoid form of $W(x)$. It can be shown that in each row of the matrix $\mathbf{B}=(b_{i,j})$, $b_{i,j}$ will decrease as $j$ increases, as compared with the case of constant $W$. This will result in a decrease of the ratio of the first to the last coefficients in the root-determining equation, and, therefore, in a decrease in the value of the modulus of the largest root of the polynomial, eq (20).

Thus, the stability conditions severely restrict us in selecting the best procedure for numerical computations of a system of simultaneous equations which replace the original integral equation of the first kind. This stability problem can be almost completely eliminated if the original integral equation of the first kind is transformed into an equation of the second kind, by differentiation. Then eq (21) becomes

$$-K(u,u)I(u)+\int_u^{N\Delta x} I(x)\frac{d}{du}K(u,x)dx=\frac{dhu}{du}. \tag{37}$$

If $K(u, u)=0$, the differentiation is repeated, until one obtained

$$\frac{d^n}{du^n}K(u,x)\Big|_{x=0}\neq 0$$

When eq (37) is replaced by a set of simultaneous linear equations, the presence of the first term in the 1.h.s. of eq (37) drastically relaxes the restrictions imposed by stability consideration: The reason is that the presence of this term increases the values of the diagonal terms in the $\mathbf{B}$ matrix. This, in turn, tends to decrease the modulus of the largest root of eq (19). For this reason, it is preferable to transform eq (2) into eq (37) whenever $\tilde{\tilde{I}}(u)$ and $K(u, x)$ are given as analytical functions. However, in our problem these functions are given as sets of points based on experimentally determined data. Their differentiations would introduce a further uncertainty in the computed results. The method of solution we developed is to be preferred only if the numerical differentiations of experimental data are to be avoided.

## 4. Perturbation Methods

One would expect that, of all difference approximations to derivatives, the ones based on central differences would lead to the most accurate results. This would be the conclusion based on formal identities between the derivative operator $\mathbf{D}^r$, defined by

$$\mathbf{D}^r I_i=\left(\frac{\mathbf{D}^r I(x)}{dx^r}\right)_{x=(i-1)\Delta x} \tag{38}$$

and the various shifting operators $\mathbf{\Delta}^r$, $\underset{\sim}{\nabla}^r$, and $\boldsymbol{\delta}^r$.

From $\mathbf{E}=e^{\mathbf{D}}$ (in units employed in this work, based on arbitrary setting of $\Delta x=1$) we have, to the first approximation,

$$\mathbf{D}^r=(\log \mathbf{E})^r\cong\mathbf{\Delta}^r\left(1-\frac{r}{2}\mathbf{\Delta}\right)\cong\boldsymbol{\nabla}^r(1+r/2\boldsymbol{\nabla})\cong\boldsymbol{\delta}^r\left(1-\frac{\boldsymbol{\delta}^2}{2.3\,!}\right). \tag{39}$$

Equation (39) shows the magnitude of the error involved in replacing an $r$-order derivative by various

$r$-order differences. It is the method, based on central differences, which leads to the standard integration formulas. For example, with $l=2$ and $p=2$, the coefficients in expansion (31) are the same as in Simpson's rule for numerical integration. With $l=p=6$, Weddle's rule is obtained. Unfortunately, the direct solution of the integral equation, based on the central-difference method, was shown to be numerically unstable. If, however, we wish to use the central difference method, we will have to employ other, indirect approaches to the algebraic solution of the integral equation. These approaches will be discussed now.

We start with

$$\sum_{j=i,\ i+1,\ \ldots}^{j=N} b_{i,j} I_j \equiv \sum_{j=i,\ i+l,\ i+2l,\ \ldots}^{j=N} I_j \int_{(j-1-l/2)\Delta x}^{(j-1+l/2)\Delta x} K[(i-1)\Delta x, x]dx$$

$$+1/2[I_{j+1}-2I_j+I_{j-1}] \int_{(j-1-l/2)\Delta x}^{(j-1+l/2)\Delta x} (x-j-1)^2 K[(i-1)\Delta x, x]dx+\ldots=h_i \cdot \qquad (40)$$

In matrix form, $\mathbf{BI}=\mathbf{h}$. Expand $\mathbf{B}=\mathbf{B}_0+\mathbf{B}_2+\mathbf{B}_4+\ldots$. The elements of the matrix $\mathbf{B}_0$, $b_0(i,j)$, are

$$b_0(i,j)=\int_{(j-1-l/2)\Delta x}^{(j-1+l/2)\Delta x} K[(i-1)\Delta x, x]dx$$

for $j=i,\ i+l,\ i+2l,\ \ldots$ and $b_0(i,j)=0$ for all other values of $j$. $\mathbf{B}_2$ is the matrix of $b_2(i,j)$, and

$$b_2(i,j)=-\int_{(j-l/2-1)\Delta x}^{(j+l/2-1)\Delta x} [x-(j-1)\Delta x]^2 K[(i-l)\Delta x, x]dx \text{ for } j=i,\ i+l,\ i+2l\ ,$$

$$b_2(i,j)=+\int_{(j-1-l/2)\Delta x}^{(j-1+l/2)\Delta x} (x-j\Delta x)^2 K[(i-1)\Delta x, x]dx$$

for $j=i+1,\ i+l+1\ \ldots\ ,$

$$b_2(i,j)=1/2\int_{(j-1-l/2)\Delta x}^{(j-1+l/2)\Delta x} [x-(j-1)\Delta x]^2 K[(i-1)\Delta x, x]dx$$

for $j=i-1$ and $b_2(i,j)=0$ for all other values of $j$. Similarly expressions for $b_4(i,j)$, etc., can be formulated. Let the zeroth-order solution be

$$\mathbf{B}_0\mathbf{I}^{(0)}=\mathbf{h}. \qquad (41)$$

Application of perturbation theory leads to the following:

$$\mathbf{B}=\mathbf{B}_0+\lambda\mathbf{B}_2+\lambda^2\mathbf{B}_4+\ldots \qquad (42)$$

$$\mathbf{I}=\mathbf{I}^{(0)}+\lambda\mathbf{I}^{(1)}+\lambda^2\mathbf{I}^{(2)}+\ldots. \qquad (43)$$

Equating equal powers of $\lambda$ in equation $\mathbf{BI}=\mathbf{h}$ leads to the following expressions for $\mathbf{I}^{(1)}$ and $\mathbf{I}^{(2)}$:

$$I^{(1)}=-\mathbf{B}_0^{-1}\,\mathbf{B}_2\mathbf{I}^{(0)} \qquad (44)$$

and

$$\mathbf{I}_1=\mathbf{I}^{(0)}+\mathbf{I}^{(1)}.$$

$$\mathbf{I}^{(2)}=-\mathbf{B}_0^{-1}\,[\mathbf{B}_4\,\mathbf{I}^{(0)}+\mathbf{B}_2\mathbf{I}^{(1)}]. \qquad (45)$$

Since the zeroth order solution by definition, satisfies the conditions for numerical stability, the elements of the matrix $\mathbf{B}_0^{-1}$ are bounded. Specifically, for $K[(i-1)\,\Delta x,\, x]=x$ and $l=2$, $\mathbf{B}_0^{-1}$ is is of the form

$$\mathbf{B}_0^{-1}=\begin{pmatrix} 1/2 & 0 & -1/2 & 0 & 0\ . \\ 0 & 1/4 & 0 & -1/4 & 0\ . \\ 0 & 0 & 1/6 & 0 & -1/6\ . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{pmatrix}$$

the matrix $\mathbf{B}_2$ is

$$\mathbf{B}_2=1/6\begin{pmatrix} -2 & 4 & -6 & 8 & -10 & . \\ 2 & -4 & 6 & -8 & 10 & . \\ 0 & 3 & -6 & 8 & -10 & . \\ 0 & 0 & 4 & -8 & 10 & . \\ 0 & 0 & 0 & 5 & -10 & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{pmatrix}$$

We obtain for the correction term, $I_i^{(1)}$, the following relation:

$$I_i^{(1)}=-1/12\,[I_{i-1}^{(0)}-2I_i^{(0)}+I_{i+1}^{(0)}]. \tag{46}$$

Thus, the correction to the zeroth order approximation is evaluated by calculating the second order differences, $\delta^2 I_i^{(0)}$ of the zeroth-order solution, and subtracting $1/12\,\delta^2 I_i^{(0)}$ from $I_i^{(0)}$.

The above scheme does not exhaust all possibilities in applying the perturbation method. One can use the zeroth-order solution based on a trapezoid rule. The $\mathbf{B}_2$ matrix will then be constructed differently than the one given in the above example. However, a typical row of the inverse matrix $\mathbf{B}^{-1}$ (again, calculated for $K[(i-1)\Delta x,\, x]\cong x$ and for $i\gg 1$) will be

$$\frac{2}{i}\,[1,-2,2,-2,\ldots]$$

resulting in even-odd oscillations in the computed values of $I_i$. This is also known as weak numerical instability. This weak instability can be greatly reduced, if certain averaging procedures are adopted. For example, one can replace, $I_i^{(1)}=\sum_{j=1}\beta^*(i,j)I_j$, where $\beta^*(i,j)$ are the elements of matrix $\mathbf{B}_0^{-1}\,\mathbf{B}_2$, by $\bar{\mathbf{I}}^{(1)}=\mathbf{DI}$, where $\delta(i,j)$, the $(i,j)$ element of the matrix $\mathbf{D}$ is given by $\delta(i,j)=$

$$\frac{1}{2}\Big[\beta^*(i,j)+\beta^*(i+1,j)\Big]. \tag{47}$$

In summary, two methods are at our disposition:

(1) A multi-step method with forward differences, referred to the point being computed and extended to the points already computed. This method requires precalculated "starting" values. This number of starting points is usually equal to the multiplicity of the method. These starting values are most conveniently computed using a single-step method.

(2) A single-step method used as zeroth-order approximation. Higher-order approximations are then computed by a perturbation method. The advantage of this procedure is in employment of

the more accurate central differences replacing the various derivatives. This method, if employed in combination with certain averaging procedures, tends to remove some of the weak instabilities (in form of even-odd oscillations) from the zeroth-order solution.

## 5. Some Computational Details

The two approaches outlined above, one based on a multi-step method and the other based on a single-step method serving as a zeroth order approximation in perturbation method, were employed in writing a general computer program (in Fortran) for the problem of numerical solution of eq (1). The first approach employs eq (34) as a starting equation for the actual computations. This equation is employed to the approximation that only terms up to $\Delta^2$ are retained in the expansion of the bracketed terms of eq (34). In the second approach the corrections to the zeroth order solution (which is given by eq (41)), were computed, by using eq (46) and also by performing averaging indicated by eq (47).

In actual computations, four different forms for the slit weighting function were used:

1. Constant weighting function (infinite slit)
2. Gaussian
3. Trapezoid, and
4. Data weighting function.

For the first three functions, parameters which describe them have to be supplied in the form of an experimentally determined set of values taken at a given constant interval. The kernel of the integral equation (given by eq (2a)) for these first three functions, is evaluated analytically. Moreover, all the required moments of the kernel (that is, integrals of the form $\int x^l K[(i-1)\Delta x, x]dx$, are also evaluated analytically. For the last assumed form for the weighting function, however, the computations of the kernel of the integral equation and its moments, have to be performed numerically. The time required to perform the entire "unsmearing" of the data is considerably longer for the case of the data-weighting function, than for either one of the other three functional forms of this function.

To our knowledge, the slit-weighting functions (1–3) are the only functions, which describe the actually measured weighting functions reasonably well, and which are also analytically tractable.

The results of the computations were subsequently tested by an inversion of the calculation procedures as follows: Using the calculated values of $\mathbf{I}$, the values of $\tilde{\tilde{\mathbf{I}}}$ were recalculated and compared with their original values. The accuracy of the numerical procedure was tested both with various trial functions for which analytical methods for solution of the original integral equations are available, and with experimental data taken from various measurements. Both approaches to the solution of the starting integral equation were tried with satisfactory results.

---

The author thanks R. J. Rubin for helpful discussions and suggestions in preparation of this publication.

## 6. References


[1] Mazur, J., and Wims, A. M., J. Res. Nat. Bur. Stand. (U.S.), **70A** (Phys. and Chem.) No. 6, 467–471 (Nov.–Dec. 1966).
[2] Shull, C. G., and Roess, L. C., J. Appl. Phys. **18,** 295 (1947).
[3] Kratky, O., and Skala, Z., Zeit. Elektrochem. **62,** 73 (1958).
[4] Kratky, O., Porod, G., and Skala, Z., Acta Physica, XIII, **96** (1960).
[5] Hendricks, R., and Schmidt, P., Acta Phys. Austriaca **26,** 97 (1967).
[6] Fox, L., An Introduction to Linear Matrix Algebra (Oxford University Press, 1965), p. 32.
[7] For a general discussion of multi-step method, see Henrici, P., Discrete Variable Methods in Ordinary Differential Equations (John Wiley & Sons, Inc., New York, 1962).
[8] Reference [7], p. 218, Theorem 5.5.
[9] Reference [7], Chapter 5.7.
[10] Hildebrand, F., Finite Difference Equations and Simulations, Chapter 2 (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968).
[11] Reference [7], p. 199.


(Paper 75B3&4–357)

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December 1971

# Publications of the National Bureau of Standards*

## Selected Abstract

Merris, R., Pierce, S., **A class of representations of the full linear group,** *J. Algebra* **17,** *No. 3, 346–351 (March 1971).*

Key words: Full linear group; irreducible character; symmetry class of tensors.

Let $V$ be a finite dimensional complex inner product space and let $\overset{m}{\otimes} V$ be the $m$th tensor space over $V$. Given a subgroup $H$ of $S_m$ and an irreducible character $\chi$ on $H$, we define a certain subspace $V_\chi^m(H)$ of $\overset{M}{\otimes} V$. If $T: V \to V$ is linear, let $\Pi \cdot T : \overset{m}{\otimes} V \to \overset{m}{\otimes} V$ be the $m$th Kronecker power of $T$. Then $V_\chi^m(H)$ is invariant under $\Pi \cdot T$ and we let $K(T)$ be the restriction of $\Pi \cdot T$ to $V_\chi^m(H)$. The map $T \to K(T)$ is a representation of the full linear group. We prove that if $\chi$ is not linear, then this representation is reducible.

Pierce, S., **Orthogonal groups of positive definite multilinear functionals,** *Pacific J. Math.* **33,** *No. 1, 183–189 (1970).*

Key words: Elementary divisors; isometry; multilinear functional.

Let $V$ be a finite dimensional vector space over the real numbers $R$ and let $T: V \to V$ be a linear transformation. If $\varphi : \overset{m}{\times} V \to R$ is a real multilinear functional and $\varphi(Tx_1, \ldots, Tx_m) = \varphi(x_1, \ldots, x_m)$, $x_1, \ldots, x_m \in V$, $^1T$ is called an isometry with respect to $\phi$. We say $\phi$ is positive definite if $\varphi(x, \ldots, x) \geq 0$ for all non-zero $X \in V$. In this paper we prove that if $\varphi$ is positive and $T$ is an isometry with respect to $\varphi$, then all eigenvalues of $T$ have modulus one and all elementary divisors of $T$ over the complex numbers are linear.

## Other NBS Publications

**J. Res. Nat. Bur. Stand. (U.S.), 75A (Phys. and Chem.), No. 3 (May–June 1971), SD Catalog No. C13.22/sec.A:75/3. $1.25.**

The solid phase photolysis and radiolysis of ethylene at 20 and 77 K. R. Gorden, Jr., and P. Ausloos.

Kinetic mass spectrometric investigation of the ion-molecule reactions occurring in $C_4$ and $C_5$ alkanes following photoionization at 106.7 and 104.8 nm. L. Wayne Sieck, S. K. Searles, and P. Ausloos.

An alumina standard reference material for resonance frequency and dynamic elastic moduli measurement. I. For use at 25 °C. R. W. Dickson and J. B. Wachtman, Jr.

Optical and mechanical properties of some neodymium-doped laser glasses. R. M. Waxler, G. W. Cleek, I. H. Malitson, M. J. Doge, and T. A. Hahn.

Effect of environment of viscous flow in inorganic oxide glasses. Joseph H. Simmons and Pedro B. Macedo.

Phase relations in the SrO-$IrO_2$-Ir system in air. C. L. McDaniel and S. J. Schneider.

The crystal structure of $BaCa(CO_3)_2$, (barytocalcite). B. Dickens and J. S. Bowen.

Ionization of hydrofluoric acid at 25 °C. P. R. Patel, E. C. Moreno, and J. M. Patel.

Vapor pressure equation for water in the range 0 to 100 °C. A. Wexler and L. Greenspan.

**J. Res. Nat. Bur. Stand. (U.S.), 75A (Phys. and Chem.), No. 4 (July–August 1971), SD Catalog No. C13.22/sec.A:75/4. $1.25.**

Inclusions in laser materials. Herbert S. Bennett.

Heat diffusion near absorbing centers in laser materials. Herbert S. Bennett.

Thermal conductivity, electrical resistivity, and thermopower of aerospace alloys from 4 to 300 K. J. G. Hust, D. H. Weitzel, and R. L. Powell.

Refractive indices of fused silica at low temperatures. R. M. Waxler and G. W. Cleek.

High-speed (subsecond) measurement of heat capacity, electrical resistivity, and thermal radiation properties of tungsten in the range 2000 to 3600 K. A. Cezairliyan and J. L. McClure.

The calculated continuous emission of a LTE hydrogen plasma. J. R. Roberts and P. A. Voigt.

The second spectrum of nickel (Ni II) (addenda and errata). A. G. Shenstone.

Tables of second virial coefficients and their first and second derivatives for the Stockmayer $(m, 6, 3)$ potential function. J. S. Gallagher and M. Klein.

**J. Res. Nat. Bur. Stand. (U.S.), 75A (Phys. and Chem.), No. 5 (September–October 1971), SD Catalog No. C13.22/sec.A: 75/5. $1.25.**

Measurement of the relative enthalpy of pure $\alpha$-$Al_2O_3$ (NBS heat capacity and enthalpy standard reference material No. 720) from 273 to 1173 K. D. A. Ditmars and T. B. Douglas.

High-precision coulometric titrations of boric acid. George Marinenko and C. E. Champion.

Determination of trace elements in ruby laser crystals by neutron activation analysis. B. A. Thompson and E. C. Miller.

Crystallography of some double sulfates and chromates. H. F. McMurdie, M. C. Morris, J. deGroot, and H. E. Swanson.

Discrimination between equations of state. J. R. Macdonald and D. R. Powell.

New absorption bands and isotopic studies of known transitions in CO. J. D. Simmons and S. D. Tilford.

A study of line shape of CO infrared emission lines. J. A. Dowling, S. Silverman, W. S. Benedict, and J. W. Quinn.

Ionization quantum yields and absorption coefficients of selected compounds at 58.4 and 73.6–74.4 nm. R. E. Rebbert and P. Ausloos.

Kinetic mass spectrometric investigation of the reactions of $t$-$C_4H_9^+$ ions with some simple polar molecules at thermal energies. Lucette Hellner and L. Wayne Sieck.

Energy levels, wave functions, dipole and quadrupole transitions of trivalent gadolinium ions in sapphire. P. H. E. Meijer.

**J. Res. Nat. Bur. Stand. (U.S.), 75C (Eng. & Instr.), No. 2, (April–June 1971), SD Catalog No. C13.22/sec. C:75/2. $1.25.**

** Publications for which a price is indicated are available by purchase from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (foreign postage, one-fourth additional). The NBS non-periodical series are also available from the National Technical Information Service (NTIS), Springfield, Va. 22151, which was formerly the Clearinghouse. Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

Heated air adiabatic saturation psychrometer. L. Greenspan.

Computation of the temperature distribution in cylindrical high-pressure furnaces. M. Waxman and J. R. Hastings.

Rotating adjustable transmission optical step attenuator. A. Cezairliyan.

An improved method for microwave power calibration, with application to the evaluation of connectors. G. F. Engen.

Apparatus for impact-fatigue testing. R. E. Schramm, R. L. Durcholz, and R. P. Reed.

A program in refractory metal thermocouple research. G. W. Burns and W. S. Hurst.

Polarization measurements as related to the corrosion of underground steel piling. W. J. Schwerdtfeger.

---

Ahearn, A. J. Quantitative analysis of solids by spark source mass spectrometry. (Proc. Intern. Conf. on Mass Spectroscopy, Kyoto, Japan, Sept. 6–13, 1969), Chapter in Recent Development in Mass Spectroscopy, K. Ogata and T. Hayakawa, Ed., pp. 150–157 (University of Tokyo Press, Tokyo, Japan, 1970).

Allpress, J. G., Roth, R. S., Structural studies by electron microscopy: Polymorphism of $ZrO_2 \cdot 12Nb_2O_5$, J. Solid State Chem. 2, 366–376 (1970).

Almer, H. E., Methods of calibrating weights for piston gages, Nat. Bur. Stand. (U.S.), Tech. Note 577, 54 pages (May 1971) 55 cents, SD Catalog No. C13.46:577.

Armstrong, G. T., Gibbs energy vs. free energy, Chem. Eng. News, Letter to the Editor, pp. 3 (April 26, 1971).

Ausloos, P., Rebbert, R. E., Sieck, L. W., Ion-molecule reactions in the radiolysis of ethane, J. Chem. Phys. **54,** No. 6, 2612–2618 (March 15, 1971).

Baker, D. V., Editor, 1970 Technical Highlights of the National Bureau of Standards, Annual Report, Fiscal Year 1970, Nat. Bur. Stand. (U.S.), Spec. Publ. 340, 258 pages (February 1971) $1.50, SD Catalog No. C13.10:340.

Baker, M. A., Study of the adherence of porcelain enamel to aluminum-use of the electron microscope and electron microprobe, Proc. Porcelain Enamel Institute Technical Forum, University of Illinois, Urbana, Ill., Oct. 7–9, 1970, 32, 48–51 (Porcelain Enamel Institute, Washington, D.C., 1971).

Ballard, D., Yakowitz, H., Spectrometer shield for stray iron x-ray radiation in the scanning electron microscope, Rev. Sci. Instr. **42,** No. 1, 14–15 (January 1971).

Ballard, D. B., Yakowitz, H., Goldstein, J. I., Study of metal in the lunar soil, Proc. 4th Annual Scanning Electron Microscope Symp., ITT Research Institute, Chicago, Ill., April 29, 1971 Part 1, 169–176 (April 29, 1971).

Becker, D. A., LaFleur, P. D., Production and certification of NBS biological standard reference materials, Proc. 4th Annual Conf. on Trace Substances in Environmental Health, June, 1970, pp. 433–435 (University of Missouri, Columbia, Mo. March 1971).

Bell, F. C., Wollin, H. F., Editors, Report of the 55th National Conference on Weights and Measures 1970, Nat. Bur. Stand. (U.S.), Spec. Publ. 342, 315 pages (April 1971) $1.75, SD Catalog No. C13.6/3:342.

Bennett, H. S., Absorbing centers in laser materials, J. Appl. Phys. **42,** No. 2, 619–630 (Feb. 1971).

Bennett, H. S., F centers in ionic crystals: Semicontinuum-polaron models and polarizable-ion models, Phys. Rev. B. **3,** No. 8, 2763–2777 (April 15, 1971).

Bennett, L. H., Swartzendruber, L. J., Interactions between small magnetic clusters in copper-rich Cu-Ni-Fe alloys, J. Appl. Phys. **42,** No. 4, 1547–1548 (March 15, 1971).

Bowman, R. R., Quantifying hazardous electromagnetic fields: Practical considerations, Nat. Bur. Stand. (U.S.), Tech. Note 389, 15 pages (April 1970) 30 cents, SD Catalog No. C13.46:389.

Boyd, M. E., Mountain, R. D., Methods for determining the second virial coefficient of a gas from speed-of-sound data, Phys. Rev. A. **2,** No. 5, 2164–2167 (November 1970).

Boyne, H. S., Laser frequency stabilization techniques and applications, Proc. 24th Annual Frequency Control Symp., on Frequency Control, Atlantic City, N.J., April 27–29, 1970, pp. 233–339 (U.S. Army Electronics Command, Fort Monmouth, N.J., 1970), IEEE Trans. Instr. Meas. **IM-20,** No. 1, 19–22 (Feb. 1971).

Boyne, H. S., Hall, J. L., Barger, R. L., Bender, P. L., Ward, J., Levine, J., Faller, J., Absolute strain measurements with a 30 meter vacuum interferometer, Proc. Conf. Laser Applications in the Geoscience, Huntington Beach, Calif., June 1969, pp. 215–225 (1970).

Breden, L. H., Technical Standards Coordinator, Fluorinated ethylene-propylene (FEP) plastic lined steel pipe and fittings, Nat. Bur. Stand. (U.S.), Prod. Stand. 34–70, 12 pages (Mar. 1971) 10 cents, SD Catalog No. C13.20/2:34–70.

Breden, L. H., Technical Standards Coordinator, Plastic heat-shrinkable film, Nat. Bur. Stand. (U.S.), Prod. Stand. 29–70. 16 pages (Jan. 1971) 15 cents, SD Catalog No. C13.20/2:29–70.

Breden, L. H., Technical Standards Coordinator, Polytetrafluoroethylene (PTFE) plastic lined steel pipe and fittings, Nat. Bur. Stand. (U.S.), Prod. Stand. 33–70, 12 pages (Mar. 1971) 10 cents, SD Catalog No. C13.20/2:33–70.

Breden, L. H., Technical Standards Coordinator, Poly(vinylidene fluoride) ($PVF_2$) plastic lined steel pipe and fittings, Nat. Bur. Stand. (U.S.), Prod. Stand. 35–70, 12 pages (Mar. 1971) 10 cents, SD Catalog No. C13.20/2:35–70.

Breden, L. H., Technical Standards Coordinator, Polystyrene plastic sheet, Nat. Bur. Stand. (U.S.), Prod. Stand. 31–70, 14 pages (January 1971) 15 cents, SD Catalog No. C13.20/2:31–70.

Brenner, A., Note on an organic-electrolyte cell with a high voltage, J. Electrochem, Soc. **118,** No. 3, 461–462 (March 1971).

Brenner, A., Anderson, H. J., Note on the preparation of rhenium oxychlorides, J. Electrochem. Soc. **118,** No. 2, 373–374 (February 1971).

Brenner, F. C., Kondo, A., Research for a uniform quality grading system for tires. IV. Tread wear, Rubber Chem. Technol. **44,** No. 1, 106–121 (March 1971).

Brower, W. S., Jr., Parker, H. S., Growth of single crystal cuprous oxide, J. Crystal Growth **8,** 227–229 (1971).

Brown, D. W., Lowry, R. E., Wall, L. A., The radiation-induced polymerization of 3,3,4,4,5,5,5-heptafluoropentene-1 at high pressure, J. Poly. Sci. **8,** Part A–1, 3483–3493 (1970).

Brown, R. L., An upper limit for the rate of destruction of $O_2(^1\Delta_g)$ by atomic hydrogen, J. Geophys. Res., Space Phys. **75,** No. 19, 3935–3936 (July 1, 1970).

Brown, R. L., Diffusion of a trace gas into a flowing carrier, Intern. J. Chem. Kinetics **2,** 475–477 (1970).

Bullis, W. M., Editor, Methods of measurement for semiconductor materials, process control, and devices, Quarterly Report July 1 to September 30, 1970, Nat. Bur. Stand. (U.S.), Tech. Note 571, 58 pages (Apr. 1971) 60 cents, SD Catalog No. C13.46:571.

Burdick, M. D., An instrumental procedure for evaluating adherence of porcelain enamel cover coats direct-to-steel, Proc. Porcelain Enamel Institute Technical Forum, University of Illinois, Urbana, Ill., Oct. 7–9, 1970, 32, 160–167 (Porcelain Enamel Institute, Washington, D.C., 1971).

Carter, G. C., Kahan, D. J., Bennett, L. H., Cuthill, J. R. Dobbyn, R. C., The NBS alloy data center: permuted materials index, Nat. Bur. Stand. (U.S.), Spec. Publ. 324, 683 pages (Mar. 1971) $7.00, SD Catalog No. C13.10:324.

Carter, R. S., Editor, Reactor Radiation Division: Annual Progress Report for the Period Ending October 31, 1970, Nat. Bur. Stand. (U.S.), Tech. Note 567, 66 pages (Mar. 1971) 70 cents, SD Catalog No. C13.46:567.

Cassidy, M. D., Calibration and test services of the National Bureau of Standards, Nat. Bur. Stand. (U.S.), Spec. Publ. 250–1970 Edition, 209 pages (Dec. 1970) $2.00, SD Catalog No. C13.10:250. Supersedes NBS Miscellaneous Publication 250–1968 Edition.

Cezairliyan, A., A high speed method of measuring thermal expansion of electrical conductors, Rev. Sci. Instr. 42, No. 4, 540–541 April 1971.

Cezairliyan, A., Short-time (msec) fluctuations of radiance temperature of graphite arc, Appl. Opt. 10, No. 5, 1178 (May 1971).

Chandler, H. H., Rupp, N. W., Paffenbarger, G. C., Poor mercury hygiene from ultrasonic amalgam condensation, J. Am. Dental Assoc. **82,** 553–557 (March 1971).

Chappell, S. E., Humphreys, J. C., Energy deposition in surface layers of silicon by monoenergetic electrons, Proc. 4th Intern. Conf. Electron and Ion Beam Science and Technology, May 10–14, 1970, Los Angeles, Calif., pp. 655–666 (The Electrochemical Society, Inc., New York, N. Y., 1970).

Chappell, S. E., Humphreys, J. C., Silicon detector measurements of energy deposition in aluminum by monoenergetic electrons

(Proc. Annual Conf. Nuclear Space Radiation, La Jolla, Calif., July 1970), IEEE Trans. Nuclear Science **NS–17,** No. 6, 272–277 (Dec. 1970).

Cohen, J. Edelman, S., Direct piezoelectric effect in polyvinylchloride films, J. Appl. Phys. **42,** No. 2, 893–894 (February 1971).

Collin, G. J., Ausloos, P., Ion-molecule reactions in the condensed-phase radiolysis of hydrocarbon mixtures. III. Reactions of $i\text{-}C_4H_8^+$ and *tert*-$C_4H_9^+$ ions originating from neopentane, J. Am. Chem. Soc. **93,** No. 6, 1336–1340 (March 24, 1971).

Colwell, J. H., Mangum, B. W., Thornton, D. D., Low-temperature magnetic susceptibility and heat capacity of $GdAsO_4$, Phys. Rev. B. **3,** No. 11, 3855–3861 (June 1, 1971).

Cooper, J. W., Conneely, M. J., Smith, K., Ormonde, S., Resonant structure of lithium between the $2^2S$ and $2^1P$ thresholds. Phys. Rev. Letters **25,** No. 22, 1540–1543 (November 30, 1970).

Costrell, L., CAMAC instrumentation system-introduction and general description (Proc. Nuclear Science Symp., New York, N.Y., Nov. 6, 1970), IEEE Trans. Nucl. Sci. **NS–18,** No. 2, 1–6 (April 1971).

Currie, L. A., Rodriguez-Pasqués, R. H., Photonuclear tritium yields at 90 MeV, Nucl. Phys. **A157,** 49–60 (1970).

Danos, M., Gibson, B. F., Very high-momentum components in nuclei and far "subthreshold" production of quarks, Phys. Rev. Letters **26,** No. 8, 473–476 (Feb. 22, 1971).

Danos, M., Gillet, V., Evidence for quartet structure in medium and heavy nuclei, Physics Letters **34B,** No. 1, 24–26 (Jan. 18, 1971).

Danos, M., Spicer, B. M., Quartet structure in light nuclei, Z. Physik **237,** 320–326 (1970).

Day, G. W., Linear and nonlinear optical properties of trigonal selenium, Appl. Phys. Letters, **18,** No. 8, 347–349 (April 15, 1971).

DeVoe, J. R., Spijkerman, J. J., Mössbauer spectrometry, Anal. Chem. Annual Reviews **42,** No. 5, 366R–388R (April 1970).

Dibeler, V. H., Photoionization studies and thermodynamic properties of some halogen molecules (Proc. Intern. Conf. on Mass Spectroscopy, Kyoto, Japan, Sept. 8–13, 1969), Chapter in Recent Development in Mass Spectroscopy, K. Ogata and T. Hayakawa, Ed., pp. 781–790 (University of Tokyo Press, Tokyo, Japan, 1970).

Dibeler, V. H., Walker, J. A., McCulloh, K. E., Photoionization study of chlorine monofluoride and the dissociation energy of fluorine, J. Chem. Phys. **53,** No. 12, 4414–4417 (December 15, 1970).

Dikkers, R. D., Marshall, R. D., Thom, H. C. S., Hurricane Camille–August 1969, Nat. Bur. Stand. (U.S.), Tech. Note 569, 71 pages (Mar. 1971) 70 cents, SD Catalog No. C13.46:569.

DiMarzio, E. A., Guttman, C. M., Separation by flow and its application to gel permeation chromatography, J. Chromatog. **55,** 83–97 (1971).

Domalski, E. S., Thermochemical properties of peroxyacetyl (PAN) and peroxybenzoyl nitrate (PBN), Environ. Sci. Tech. Communications, **5,** 443–444 (1971).

Dziuba, R. F., Dunfee, B. L., Resistive voltage-ratio standard and measuring circuit, IEEE Trans. Instr. Meas. **IM–19,** No. 4, 266–277 (Nov. 1970).

Edgerly, D. E., Weights and Measures labeling Handbook, Nat. Bur. Stand. (U.S.), Handb. 108, 406 pages (May 1971) $3.25, SD Catalog No. C13.11:108.

Edqvist, O., Lindholm, E., Selin, L. E., Asbrink, L., Kuyatt, C. E., Mielczarek, S. R., Simpson, J. A., Fischer-Hjalmars, I., Rydberg series of small molecules VIII photoelectron spectroscopy and electron spectroscopy of $NO_2$, Phys. Script. **1,** 172–178 (1970).

Engen, G. F., A new method of characterizing amplifier noise performance, IEEE Trans. Instr. Meas. **IM–19,** No. 4, 344–349 (Nov. 1970).

Evans, B. J., Swartzendruber, L. J., Mössbauer study of local molecular fields, cation distributions, and recoilless fractions in Sb-substituted lithium ferrite, J. Appl. Phys. **42,** No. 4, 1628–1630 (March 15, 1971).

Evenson, K. M., Radford, H. E., Moran, M. M., Jr., CH free radicals detected by infrared laser magnetic resonance, appl. Phys. Letters **18,** No. 10, 426–429 (May 15, 1971).

Evenson, K. M., Wells, J. S., Matarrese, L. M., Jennings, D. A., Variable output-coupling far-infrared Michelson laser, J. Appl. Phys. **42,** No. 3, 1233–1234 (March 1, 1971).

Farabaugh, E. N., Mauer, F. A., Application of a modified technique for x-ray topography of large crystals, J. Crystal Growth **7,** 282–284 (1970).

Fatiadi, A. J., Determination of inososes with an alkaline solution of copper (II) oxalate-tartrate complex (the Somogyi reagent) and reaction mechanisms involved, Carbohydrate Res. **17,** 419–430 (March 1971).

Fatiadi, A. J., Oxidative cleavage of aromatic azines with periodic acid: e.s.r. evidence for the liberation of nitrogen dioxide, Chem. Ind. pp. 64–66 (Jan. 9, 1971).

Franklin, A. D., Bennett, H. S., Editors, ARPA–NBS Program of research on high temperature materials and laser materials, January 1 to June 30, 1970, Nat. Bur. Stand. (U.S.), Tech. Note 565, 84 pages (Jan. 1971) 75 cents, SD Catalog No. C13.46:565.

Fickett, F. R., Resistivity of polycrystalline aluminum and copper in high magnetic fields: The effect of temperature and purity, Appl. Phys. Letters **17,** No. 12, 525–527 (Dec. 15, 1970).

Fickett, F. R., Clark, A. F., Longitudinal magnetoresistance anomalies, J. Appl. Phys. **42,** No. 1, 217–219 (Jan. 1971).

Franklin, A. D., Bennett, H. S., Editors, ARPA–NBS program of research on high temperature materials and laser materials, Reporting Period July 1 to December 31, 1970, Nat. Bur. Stand. (U.S.), Tech. Note 574, 44 pages (May 1971) 50 cents, SD Catalog No. C13.46:574.

Freeman, D. H., Zielinski, W. L., Jr., Editors, Separation and purification section: Summary of activities, July 1969 to June 1970, Nat. Bur. Stand. (U.S.), Tech. Note 549, 86 pages (Jan. 1971) 75 cents, SD Catalog No. C13.46:549.

Furcolow, W. H., Technical Standards Coordinator, Clinical thermometers, (Maximum-self-registering, mercury-in-glass), Nat. Bur. Stand. (U.S.), Prod. Stand. 39–70, 12 pages (May 1971) 15 cents, SD Catalog No. C13.20/2:39–70.

Furcolow, W. H., Technical Standards Coordinator, Hinged interior wood door units, Nat. Bur. Stand. (U.S.), Prod. Stand. 32–70, 15 pages (Mar. 1971) 30 cents, SD Catalog No. C13.20/2:32–70.

Gadzuk, J. W., Plummer, E. W., Energy distributions for thermal field emission, Phys. Rev. B. **3,** No. 7, 2125–2129 (April 1, 1971).

Gadzuk, J. W., Plummer, E. W., Hot-hole-electron cascades in field emission from metals, Phys. Rev. Letters **26,** No. 2, 92–95 (Jan. 11, 1971).

Garner, E. L., Machlan, L. A., Shields, W. R., Standard Reference Materials: Uranium isotopic standard reference materials, (Certification of uranium isotopic standard reference materials), Nat. Bur. Stand. (U.S.), Spec. Publ. 260–27, 162 pages (April 1971) $1.25, SD Catalog No. C13.10:260–27.

Garvin, D., The information analysis center and the library, Spec. Libraries **62,** No. 1, 17–23 (January 1971).

Gebbie, K. B., Thomas, R. N., On the dependence of $T_e$ upon quantity versus quality of the radiation field in a stellar atmosphere, Astrophys. J. **161,** No. 1, 229–241 (July 1970).

Geltman, S., Burke, P. G., Electron scattering by atomic hydrogen using a pseudo-state expansion II. Excitation of 2*s* and 2*p* states near threshold, J. Phys. B: Atom. Molec. Phys. **3,** No. 8, 1062–1072 (August 1970).

Geltman, S., Burke, P. G., Electron scattering by atomic hydrogen using a pseudo-state expansion II. Excitation of 2*s* and 2*p* states near threshold, J. Phys. B: Atom. Molec. Phys. **3,** No. 8, 1062–1072 (Aug. 1970).

Godfrey, J. T., Vidal, C. R., Smith, E. W., Cooper, J., Effect of time ordering in the unified theory, Phys. Rev. A. **3,** No. 5, 1543–1546 (May 1971).

Green, M. S., Cooper, M. J., Sengers, J. M. H. L., Extended thermodynamic scaling from a generalized parametric form, Phys. Rev. Letters **26,** No. 9, 492–495 (March 1, 1971).

Greenspan, M., Comment on "Response of ideal circuit elements to sinusoidal excitations", Am. J. Phys. **39,** No. 5, 586–587 (May 1971).

Harman, G. G., Kessler, H. K., Application of capacitor microphones and magnetic pickups to the tuning and trouble shooting of microelectronic ultrasonic bonding equipment, Nat. Bur. Stand. (U.S.), Tech. Note 573, 24 pages (May 1971) 35 cents, SD Catalog No. C13.46:573.

Harris, F. K., Fowler, H. A., Olsen, P. T., Accurate Hamon-pair potentiometer for Josephson frequency-to-voltage measurements, Metrologia **6,** No. 4, 134–142 (October 1970).

Hastie, J. W., Hauge, R. H., Margrave, J. L., Infrared spectra and geometries of heavy metal halides: $SrCl_2$, $BaCl_2$, $EuCl_2$, $EuF_2$, $PbCl_2$, and $UCl_2$, High Temp. Sci. **3,** No. 1, 56–71 (Jan. 1971).

Hellwig, H., Areas of promise for the development of future primary frequency standards, Proc. 24th Annual Symp. on Frequency Control, Atlantic City, N.J., April 27–29, 1970, pp. 246–258 (Electronic

Industries Assn., Washington, D.C., 1970).
Hellwig, H., Areas of promise for the development of future primary frequency standards, Metrologia **6,** No. 4, 118–126 (October 1970).
Hellwig, H., Vessot, R. F. C., Levine, M. W., Zitzewitz, P. W., Allan, D. W., Glaze, D. J., Measurement of the unperturbed hydrogen hyperfine transition frequency, IEEE Trans. Instr. Meas. **IM–19,** No. 4, 200–209 (November 1970).
Herron, J. T., Huie, R. E., Mass spectrometric studies of the reactions of singlet oxygen in the gas phase, Ann. N.Y. Acad. Sci. 171, Article **1,** 229–238 (Oct. 15, 1970).
Hicho, G. E., Yakowitz, H., Rasberry, S. D., Michaelis, R. E., Standard reference materials: A standard reference material containing, nominally four percent austenite, Nat. Bur. Stand. (U.S.), Spec. Publ. 260–25, 22 pages (Feb. 1971) 30 cents, SD Catalog No. C13.10:260–25.
Hoer, C. A., Agy, D. L., Broad-band resistive-divider-type directional coupler, IEEE Trans. Instr. Meas. **IM–19,** No. 4, 336–343 (November 1970).
Hord, J., Voth, R. O., Tabulated values of cavitation *B*-factor for helium, $H_2$, $N_2$, $F_2$, $O_2$, refrigerant 114, and $H_2O$, Nat. Bur. Stand. (U.S.), Tech. Note 397, 116 pages (Feb. 1971) $1.00, SD Catalog No. C13.46:397.
Hudson, R. D., Kieffer, L. J., Compilation of atomic ultraviolet photoabsorption cross sections for wavelengths between 3000 and 10 Å, Atomic Data **2,** No. 3, 205–262 (May 1971).
Hummer, D. G., Rybicki, G. B., Non-coherent scattering—VII. Frequency-dependent thermalization lengths and scattering with continuous absorption, Monthly Notices Roy. Astron. Soc. **150,** No. 4, 419–434 (1970).
Hussmann, E. K., A holographic interferometer for measuring radiation energy deposition profiles in transparent liquids, Appl. Opt. **10,** No. 1, 182–186 (January 1971).
Hustvedt, E. H., Conference Coordinator, Friend, M. L., Editor, Proceedings of joint meeting of government operations research users and producers, Held at the National Bureau of Standards, Gaithersburg, Maryland, June 5–6, 1969, Nat. Bur. Stand. (U.S.), Spec. Publ. 347, 164 pages (May 1971) $1.25, SD Catalog No. C13.10:347.
Ives, L. K., Ruff, A. W., Jr., Extended dislocation configurations in HCP silver-tin alloys of low stacking-fault energy, Metal Sci. J. **4,** 201–209 (May 1970).
Jacox, M. E., Milligan, D. E., Infrared spectrum and structure of the species $CO_3$, J. Chem. Phys. **54,** No. 3, 919–926 (February 1, 1971).
Jennings, D. A., Power and energy measurement of repetitively pulsed lasers, Nat. Bur. Stand. (U.S.), Tech. Note 398, 10 pages (March 1971) 25 cents, SD Catalog No. C13.46:398.
Julienne, P. S., Predissociation of the $H_2$ $D^1\Pi_u$ state, Chem. Phys. Letters **8,** No. 1, 27–28 (January 1, 1971).
Kamas, G., Hanson, D. W., Recent VHF/UHF satellite timing experiments at the National Bureau of Standards, Proc. Precise Time and Time Interval (PTTI) Strategic Planning Meeting (U), Washington, D.C., December 10–11, 1970, 1, 79–82 (1970).
Kamper, R. A., Zimmerman, J. E., Noise thermometry with the Josephson effect, J. Appl. Phys. **42,** No. 1, 132–136 (Jan. 1971).
Keysar, B. C., Specifications and tolerances for reference standards and field standard weights and measures. 2. Specifications and tolerances for field standard measuring flasks, Nat. Bur. Stand. (U.S.), Handb. 105–2, 6 pages (Jan. 1971) 25 cents, SD Catalog No. C13.11:105–2.
Keysar, B. C., Specifications and tolerances for reference standards and field standard weights and measures. 3. Specifications and tolerances for metal volumetric field standards, Nat. Bur. Stand. (U.S.), Handb. 105–3, 8 pages (May 1971) 25 cents, SD Catalog No C13.11:105–3.
Kidnay, A. J., Hiza, M. J., The purification of helium by physical adsorption at 76 °K, AIChE J. **16,** No. 7, 949–954 (Nov. 1970).
Kieffer, L. J., Low energy electron collision cross section data. Part II. Electronic excitation level and line cross sections, Atomic Data **1,** No. 2, 121–287 (Nov. 1969).
Klein, M., Hanley, H. J. M., The *m*-6-8 potential function, J. Chem. Phys. **53,** No. 12, 4722–4723 (Dec. 15, 1970).
Klose, J. Z., Mean lives and absolute *f*-values in neutral iron, Astrophys. J. **165,** No. 3, 637–642 (May 1, 1971).
Kruger, J., Frankenthal, R. P., B. Corrosion of metals, Chapter in Volume **IV,** Techniques in Metals Research, Chapter 10B oxidation and corrosion, R. A. Rapp, Ed., R. F. Bunshah, Series Ed., 14, Part 2, 571–667 (John Wiley and Sons Inc., New York, N.Y. 1970).
Kulin, G., Gurewitz, P. H., Editors, Hydraulic Research in the United States 1970—Including Contributions from Canadian Laboratories, Nat. Bur. Stand. (U.S.), Spec. Publ. 346, 354 pages (Mar. 1971) $2.50, SD Catalog No. C13.10:346.
Kuriyama, M., Miyakawa, T., Primary and secondary extinctions in the dynamical theory for an imperfect crystal, Acta Cryst. **A26,** Part 6, 667–673 (November 1970).
Kurylo, M. J., Peterson, N. C., Braun, W., Absolute rate of the reaction $H + H_2S$, J. Chem. Phys. **54,** No. 3, 943–946 (February 1, 1971).
Kurylo, M. J., Peterson, N. C., Braun, W., Temperature and pressure effects in the addition of H atoms to propylene, J. Chem. Phys. **54,** No. 11, 4662–4666 (June 1, 1971).
Lane, N. F., Geltman, S., Differential elastic and rotational excitation cross sections for electron-$H_2$ scattering, Phys. Rev. **184,** No. 1, 46–51 (Aug. 5, 1969).
Latanision, R. M., Ruff, A. W., The temperature dependence of stacking fault energy in Fe-Cr-Ni alloys, Met. Trans. **2,** 505–509 (February 1971).
Laufer, A. H., Low-temperature chromatographic determination of ketene and methyl ketene, J. Chromatogr. Sci. **8,** 677–678 (Nov. 1970).
Laufer, A. H., Keller, R. A., Lowest excited states of ketene, J. Am. Chem. Soc. **93,** No. 1, 61–63 (January 1971).
Linsky, J. L., Avrett, E. H., The solar H and K lines, Astron. Soc. Pacific **82,** 169–248 (April 1970).
Liu, Y. M., Coleman, J. A., Electron radiation damage effects in silicon surface-barrier detectors, IEEE Trans. Nucl. Sci. **NS–18,** No. 1, 192–199 (February 1971).
Lloyd, E. C., Editor, Accurate characterization of the high-pressure environment, Proceedings of a Symposium held at the National Bureau of Standards, Gaithersburg, Maryland, October 14–18, 1968, Nat. Bur. Stand. (U.S.), Spec. Publ. 326, 343 pages (March 1971) $4.50, SD Catalog No. C13.10:326.
Lutz, G. J., Photon activation analysis—A Review, Anal. Chem. **43,** No. 1, 93–104 (January 1971).
McClellan, B. E., Menis, O., Kinetics and mechanism of extraction of iron (III) with β-isopropyltropolone, Anal. Chem. **43,** No. 3, 436–438 (March 1971).
McConnell, P. M., Daney, D. E., Kirgis, J. B., Thermoelastic expansion and creep of polyethylene terephthalate and polypyromelitimide film and polyethylene terephthalate fibers from 20 to 295 K, J. Appl. Phys. **41,** No. 13, 5066–5070 (Dec. 1970).
McDonald, D. G., Evenson, K. M., Wells, J. S., Cupp, J. D., High-frequency limit of the Josephson effect, J. Appl. Phys. **42,** No. 1, 179–181 (Jan. 1971).
McLaughlin, W. L., Hussmann, E. K., Eisenlohr, H. H., Chalkley, L., A chemical dosimeter for monitoring gamma-radiation doses of 1–100 krad, Intern. J. Appl. Radiation Isotopes **22,** 135–140 (1971).
Mackay, D. R., Technical Standards Coordinator, Editorial format for Voluntary Product Standards. Part 1. Contents of a Voluntary Product Standard, Nat. Bur. Stand. (U.S.), Prod. Stand. 0–70, 13 pages (March 1971) 15 cents, SD Catalog No. C13.20/2:0–70. Supersedes PS 0–67.
Madey, T. E., Yates, J. T., Jr., Electron-stimulated desorption and work function studies of clean and cesiated (110) GaAs, J. Vacuum Sci. Technol. **8,** No. 1, 39–44 (January–February 1971).
Mandel, J., A new analysis of variance model for non-additive data, Technometrics **13,** No. 1, 1–18 (February 1971).
Manning, J. R., Correlation effects and activation energies for diffusion in alloys (Proc. Conf. Atomic Transport in Solids and Liquids, Marstand, Sweden, June 1970), Z. Naturforsch. **26A,** No. 1, 69–76 (Jan. 1971).
Martin, J. F., Standard Reference Materials: National Bureau of Standards-U.S. Steel corporation joint program for determining oxygen and nitrogen in steel, Nat. Bur. Stand. (U.S.), Spec. Publ. 260–26, 40 pages (February 1971), 50 cents, SD Catalog No. C13.10:260–26.
Mason, H. L., Editor, Innovative metrology—key to progress, Proceedings of the 1970 Standards Laboratory Conference, Nat. Bur. Stand. (U.S.), Spec. Publ. 335, 132 pages (Mar. 1971) $1.50, SD Catalog No. C13.10:335.
Matarrese, L. M., Evenson, K. M., Improved coupling to infrared whisker diodes by use of antenna theory, Appl. Phys. Letters **17,** No. 1, 8–10 (July 1, 1970).
Meijer, P. H. E., Influence of the dipole-dipole coupling on the specific heat of cesium titanium alum, Phys. Rev. **B3,** No. 1, 182–185 (January 1, 1971).

Meinke, W. W., Standard reference materials for clinical measurements, Anal. Chem. **43,** 28A–47A (May 1971).

Menis, O., McClellan, B. E., Bright, D. S., Determination of the formation constants of iron (III) and vanadium (V) with β-isopropyltropolone using the extraction method, Anal. Chem. **43,** No. 3, 431–435 (March 1971).

Menis, O., Rains, T. C., Sensitivity, detection limit, precision and accuracy in flame emission and atomic absorption spectrometry, Chapter 2, Analytical Flame Spectroscopy, Selected Topics, R. Mavrodineanu, Ed., pp. 47–77 (Springer-Verlag Publishing Co., New York, N.Y., 1970).

Milligan, D. E., Jacox, M. E., Infrared spectrum and structure of intermediates in the reaction of OH with CO, J. Chem. Phys. **54,** No. 3, 927–942 (February 1, 1971).

Missoni, G., Dick, C. E., Placious, R. C., Motz, J. W., Inelastic atomic scattering of 0.1-, 0.2-, 0.4- and 3.0-MeV electrons, Phys. Rev. A. **2,** No. 6, 2309–2317 (December 1970).

Mohan, R., Danos, M., Biedenharn, L. C., Isospin impurities in the nuclear ground states, Phys. Rev. C. **3,** No. 2, 468–479 (Feb. 1971).

Moore, C. E., Annual report on spectroscopy, July 1, 1969–June 30, 1970, Bull. Am. Astron. Soc. **3,** No. 1, 154–155 (1971).

Motz, J. W., Dick, C. E., Lucas, A. C., Placious, R. C., Sparrow, J. H., Production of high-intensity K x-ray beams, J. Appl. Phys. **42,** No. 5, 2131–2133 (April 1971).

Mountain, R. D., Comment on "Rotational diffusion of spherical-top molecules in liquids," J. Chem. Phys. **54,** No. 7, 3243 (April 1, 1971).

Mukerjee, P., Mysels, K. J., Critical micelle concentrations of aqueous surfactant systems, Nat. Stand. Ref. Data Ser., Nat. Bur. Stand. (U.S.), 36, 227 pages (Feb. 1971) $3.75, SD Catalog No. C13.48:36.

Neill, A. H., Jr., Wet traction of tractionized tires, Nat. Bur. Stand. (U.S.), Tech. Note 566, 14 pages (Feb. 1971) 20 cents, SD Catalog No. C13.46:566.

Nimeroff, I., Color-match classification by variable parameters, Color. Engr. **9,** No. 2, 13–17 (March–April 1971).

Noack, M., Kokoszka, G. F., Gordon, G., Dynamic Jahn-Teller effects and magnetic anisotropies in aqueous solutions and water-ethanol glasses of copper (II) solvates and complexes with 2,2′-dipyridine, J. Chem. Phys. **54,** No. 3, 1342–1350 (Feb. 1, 1971).

Ogburn, F., Shives, T. R., Brady, C. H., Calibration of a filar micrometer, J. Mater. **6,** No. 1, 60–66 (March 1971).

Olson, W. B., Papousek, D., High-resolution infrared spectra of ethanelike molecules and the barrier to internal rotation: The $\nu_9$, $\nu_{13}$ band of dimethylacetylene, J. Mol. Spectry. **37,** No. 3, 527–534 (March 1971).

Osinsky, V. I., Winogradoff, N. N., Excitation and temperature dependence of band-edge photoluminescence in gallium arsenide, Phys. Rev. B. **3,** No. 10, 3341–3346 (May 15, 1971).

Page, C. H., Comments on "Nonlinear resistors that generate subharmonics", Proc. IEEE Letters **59,** No. 2, 301 (Feb. 1971).

Page, C. H., Vigoureux, P., Editors, The international system of units (SI), Nat. Bur. Stand. (U.S.), Spec. Publ. 330, 42 pages (Jan. 1971), 50 cents, SD Catalog No. C13.10:330.

Parks, E. J., Thermal analysis of modified cellulose, Tappi **54,** No. 4, 537–544 (April 1971).

Paule, R. C., Mandel, J., Analysis of interlaboratory measurements on the vapor pressures of cadmium and silver (Certification of Standard Reference Materials 746 and 748), Nat. Bur. Stand. (U.S.), Spec. Publ. 260–21, 30 pages (Jan. 1971), 35 cents, SD Catalog No. C13.10:260–21.

Peavy, S. T., Varner, R. N., Hogben, D., Source listing of OMNITAB II program, Nat. Bur. Stand. (U.S.), Spec. Publ. 339, 371 pages (Dec. 1970) $4.75, SD Catalog No. C13.10:339.

Pecker, J. C., Thomas, R. N., Personal libraries and publication policies, Atoms **25,** 37–38 (Jan. 1970).

Peterson, R. L., Numerical study of deformation-potential scattering of electrons by optical phonons in a longitudinal magnetic field, Phys. Rev. B. **2,** No. 10, 4135–4144 (November 15, 1970).

Phucas, C. B., Technical Standards Coordinator, Package quantities of green olives, Nat. Bur. Stand. (U.S.), Prod. Stand. 40–70, 9 pages (May 1971) 10 cents, SD Catalog No. C13.20/2:40–70.

Phucas, C. B., Technical Standards Coordinator, Package quantities of instant nonfat dry milk, Nat. Bur. Stand. (U.S.), Prod. Stand. 37–70, 9 pages (April 1971) 10 cents, SD Catalog No. C13.20/2:37–70.

Phucas, C. B., Technical Standards Coordinator, Package quantities of instant mashed potatoes, Nat. Bur. Stand. (U.S.), Prod. Stand. 41–70, 8 pages (April 1971) 10 cents, SD Catalog No. C13.20/2:41–70.

Ramboz, J. D., A link-compensated ratio transformer bridge, Rev. Sci. Instr. **42,** No. 4, 522–524 (April 1971).

Redding, R. W., Hougen, J. T., Calculation of the Eulerian angles and large amplitude vibrational coordinate corresponding to an arbitrary instantaneous molecular configuration, J. Mol. Spectry. **37,** No. 2, 366–370 (February 1971).

Reitwiesner, G. W., On computer performance measurement programming measuring indexing adroitness by isolating complex primes, Nat. Bur. Stand. (U.S.), Tech. Note 572, 25 pages (Apr. 1971) 35 cents, SD Catalog No. C13.46:572.

Richmond, J. C., Editor, Space simulation. Proceedings of a Conference held at NBS, Gaithersburg, Maryland, September 14–16, 1970. Nat. Bur. Stand. (U.S.), Spec. Publ. 336, 984 pages (October 1970), $5.25, SD Catalog No. C13.10:336.

Risley, A. S., The physical basis of atomic frequency standards, Nat. Bur. Stand. (U.S.), Tech. Note 399, 54 pages (April 1971) 60 cents, SD Catalog No. C13.46:399.

Robertson, B., Mitchell, W. C., Equations of motion in nonequilibrium statistical mechanics. III. Open systems, J. Math. Phys. **12,** No. 3, 563–568 (March 1971).

Robinson, D. C., Determining of load errors in universal testing machines due to speed of testing, Mater. Res. Stand. **11,** No. 1, 14–17 (January 1971).

Rosenstock, H. M., Botter, R., Franck-Condon principle for the ionization of polyatomic molecules, (Proc. Intern. Conf. on Mass Spectroscopy, Kyoto, Japan, Sept. 1969), Chapter in Recent Development in Mass Spectroscopy, K. Ogata and T. Hayakawa, Ed., pp. 797–806 (University of Tokyo Press, Tokyo, Japan, 1970).

Rowland, G. A., The increasing role of states in building codes and standards, Proc. 26th. Annual Conf. Reinforced Plastics, Composites Division, Washington, D.C., February 9–12, 1971, Section 5–B, pp. 1–3 (February 10, 1971).

Ruegg, F. C., A circuit which expands the count storage capacity of a multichannel analyzer, Nucl. Instr. Methods **92,** 7–11, (1971).

Scharf, K., Spectrophotometric measurement of ferric ion concentration in the ferrous sulphate (Fricke) dosimeter, Phys. Med. Biol. **16,** No. 1, 77–86 (1971).

Schramm, R. E., Corrections and calculations on a x-ray diffraction line profile: A computer program, Nat. Bur. Stand. (U.S.), Tech. Note 600, 34 pages (June 1971) 40 cents, SD Catalog No. C13.46:600.

Schroeder, L. W., Rush, J. J., Neutron diffraction study of the structure and thermal motion of phosphonium bromide, J. Chem. Phys. **54,** No. 5, 1968–1973 (March 1, 1971).

Scott, W. W., Jr., New coaxial RF-DC ammeter, IEEE Trans. Instr. Meas. **IM–19,** No. 4, 318–323 (Nov. 1970).

Sieck, L. W., Searles, S., Ausloos, P., High-pressure photoionization mass spectrometry. Photoionization of propane at 11.6–11.8 eV. Formation and reactivity of the $(C_3H_8)_2^+$ dimer ion, J. Chem. Phys. **54,** No. 1, 91–95 (Jan. 1, 1971).

Simmons, J. H., Macedo, P. B., Analysis of viscous relaxation in critical oxide mixtures, J. Chem. Phys. **54,** No. 3, 1325–1331 (February 1, 1971).

Sixsmith, H., Giarratano, P., A miniature centrifugal pump, Rev. Sci. Instr. **41,** No. 11, 1570–1573 (Nov. 1970).

Slattery, W. J., Editor, An index of U.S. voluntary engineering standards, Nat. Bur. Stand. (U.S.), Spec. Publ. 329, 1,000 pages (Mar. 1971) $9.00, SD Catalog No. C13.10:329.

Smith, C. E., Milligan, D. E., Jacox, M. E., Matrix-isolation study of the vacuum-ultraviolet photolysis of chlorofluoromethane. The infrared and ultraviolet spectra of the free radical ClCF, J. Chem. Phys. **54,** No. 7, 2780–2793 (April 1, 1971).

Smith, R. W., Jr., Conference Chairman, Editor, Precoordination-basis for industrialized building, (Proc. Conf. Gaithersburg, Md., Sept. 24–26, 1969), Nat. Bur. Stand. (U.S.), Bldg. Sci. Ser. 32, 136 pages (Jan. 1971) $1.50, SD Catalog No. C13.29/2:32.

Somes, N. F., Dikkers, D. R., Boone, T. H., Lubbock Tornado: A survey of building damage in an urban area, Nat. Bur. Stand. (U.S.), Tech. Note 558, 38 pages (Mar. 1971) 50 cents, SD Catalog No. C13.46:558.

Spiegel, V., Jr., Murphey, W. M., Calculation of thermal neutron absorption in cylindrical and spherical neutron sources, Metrologia, **7,** No. 1, 34–38 (Jan. 1971).

Spiegel, V., Jr., Murphey, W. M., Computer code for the calculation of thermal neutron absorption in spherical and cylindrical neutron

sources, Nat. Bur. Stand. (U.S.), Tech. Note 576, 25 pages (May 1971) 35 cents, SD Catalog No. C13.46:576.

Spijkerman, J. J., Travis, J. C., Pella, P. A., DeVoe, J. R., Preliminary study on the characteristics and design parameters for a Mössbauer resonant detector, Nat. Bur. Stand. (U.S.), Tech. öössbauer resonant detector, Nat. Bur. Stand. (U.S.), Tech. Note 541, 65 pages (Jan. 1971), 65 cents, SD Catalog No. C13.46: 541.

Sullivan, D. B., Peterson, R. L., Kose, V. E., Zimmerman, J. E., Generation of harmonics and subharmonics of the Josephson oscillation, J. Appl. Phys. **41,** No. 12, 4865–4873 (November 1970).

Swartzendruber, L. J., Localized moments on Fe impurities in Nb-Mo alloys: Mössbauer-effect absorber study, J. Appl. Phys. **42,** No. 4, 1 (March 15, 1971).

Tate, E. L., Improving library collections by weeding, Proc. U.S. Department of Interior, 1970 Departmental Library Workshop, Washington, D.C., Sept. 28–Oct. 2, 1970, pp. 54–61 (U.S. Department of the Interior Office of Library Services, Washington, D.C., Feb. 1971).

Taylor, B. N., Report on the International Conference on Precision Measurement and Fundamental Constants, held at the National Bureau of Standards, Gaithersburg, Md., Aug. 3–7, 1970, Metrologia **7,** No. 1, 39–42 (Jan. 1971).

Tech, J. L., A high-dispersion spectral analysis of the Ba II Star HD 204075 (ζ Capricorni), Nat. Bur. Stand. (U.S.), Monogr. 119, 174 pages (March 1971) $3.25, SD Catalog No. C13.44:119.

Thurber, W. R., Determination of deep impurities in silicon and germanium by infrared photoconductivity, Nat. Bur. Stand. (U.S.), Tech. Note 570, 13 pages (Mar. 1971) 25 cents, SD Catalog No. C13.46:570.

Tighe, N. J., Microstructure of fine-grain ceramics, (Proc. 15th Sagamore Army Materials Research Conf., Sagamore Conference Center, Raquette Lake, N.Y., Aug. 20–23, 1968), Chapter in Ultrafine-Grain Ceramics, pp. 109–133 (Syracuse University Press, Syracuse, N.Y., 1970).

Tipson, R. S., Brady, R. F., Jr., West, B. F., Cyclic acetals of ketoses. Part IV. Re-investigation of the oxidation of 1,2:4,5-Di-O-isopropylidene-β-D-fructopyranose with methyl sulfoxide-acetic anhydride, Carbohydrate Res. **16,** 383–393 (March 1971).

Uzgiris, E. E., Hall, J. L., Barger, R. L., Precision infrared Zeeman spectra of $CH_4$ studied by laser-saturated absorption, Phys. Rev. Letters **26,** No. 6, 289–293 (Feb. 8, 1971).

Van Blerkom, D., Hummer, D. G., The normalized on-the-spot approximation for line transfer problems, J. Quant. Spectrosc. Radiat. Transfer Note **9,** No. 11, 1567–1571 (Nov. 1969).

Varner, R. N., Peavy, S. T., Test problems and results for OMNITAB II, Nat. Bur. Stand. (U.S.), Tech. Note 551, 190 pages (Dec. 1970) $1.50, SD Catalog No. C13.46:551.

Vidal, C. R., Cooper, J., Smith, E. W., Hydrogen stark broadening calculations with the unified classical path theory, J. Quant. Spectry. Radiative Transfer **10,** No. 9, 1011–1063 (September 1970).

Vidal, C. R., Cooper, J., Smith, E. W., Unified theory calculations of stark broadened hydrogen lines including lower state interactions, Nat. Bur. Stand. (U.S.), Mongr. 120, 45 pages (Jan. 1971), 50 cents, SD Catalog No. C13.44:120.

Vidal, C. R., Cooper, J., Smith, E. W., Unified theory calculations of stark broadened hydrogen lines including lower state interactions (Short Version), J. Quant. Spectry. Radiative Transfer **11,** 263–281 (1970).

Wachtman, J. B., Jr., Standard materials for measurements on ceramics, Am. Ceram. Soc. Bull. **50,** No. 3, 242–247 (March 1971).

Wagman, D. D., Evans, W. H., Parker, V. B., Halow, I., Bailey, S. M., Schumm, R. H., Churney, K. L., Selected values of chemical thermodynamic properties—Tables for elements 54 through 61 in the Standard order of arrangements, Nat. Bur. Stand. (U.S.), Tech. Note 270–5. 49 pages (Mar. 1971) 55 cents, SD Catalog No. C13.46:270–5.

Wait, D. F., The precision measurement of noise temperature of mismatched noise generators, IEEE Trans. Microwave Theory Tech. **MTT-18,** No. 10, 715–724 (Oct. 1970).

Wallace, B., Brown, W. E., Stoichiometric composition of whitlockite, J. Dental Res. **50,** No. 2, 343–346 (March–April 1971).

Walther, H., Hall, J. L., Tunable dye laser with narrow spectral output, Appl. Phys. Letters **17,** No. 6, 239–242 (September 15, 1970).

Wampler, R. H., A report on the accuracy of some widely used least squares computer programs, J. Am. Stat. Assoc. **65,** No. 330, 549–565 (June 1970).

Watson, R. E., Bennett, L. H., Carter, G. C., Weisman, I. D., Comments on the Knight shift in bismuth and other *p*-band diamagnetic metals, Phys. Rev. B. **3,** No. 1, 222–225 (January 1971).

Waxler, R. M., Laser glass composition and the possibility of eliminating electrostrictive effects, IEEE, J. Quantum Electron. **QE-7,** No. 4, 166–167 (April 1971).

Weber, L. A., Density and compressibility of oxygen in the critical region, Phys. Rev. A. **2,** No. 6, 2379–2388 (Dec. 1970).

Weber, L. A., Some vapor pressure and *P, V, T* data on nitrogen in the range 65 to 140 K. J. Chem. Thermodynamics **2,** No. 6, 839–846 (Nov. 1970).

Weir, C. E., Piermarini, G. J., Block, S., On the crystal structures of Cs II and Ga II, J. Chem. Phys. **54,** No. 6, 2768–2770 (March 15, 1971).

Weiss, A. W., A review of theoretical developments in atomic *f*-values, (Proc. 2nd. Intern. Conf. Beam-Foil Spectroscopy in Nuclear Instruments and Methods, Lysekil, Sweden, June 8–12, 1970), Part II, Lifetimes and transition probabilities, 25–119 in Nucl. Instr. Methods **90,** 121–131 (North-Holland Publ. Co., Amsterdam, The Netherlands, 1970).

Weiss, A. W., Symmetry-adapted pair correlations in Ne, $F^-$, $Ne^+$, and F, Phys. Rev. A. **3,** No. 1, 126–129 (January 1971).

Weiss, B. Z., Meyerson, M. R., Fatigue crack initiation and propagation in chromium diffusion coated Ti bearing steel, J. Iron Steel Inst. **208,** No. 12, 1069–1077 (December 1970).

Weiss, M. J., Kuyatt, C. E., Mielczarek, S., Inelastic electron scattering from formaldehyde, J. Chem. Phys. **54,** No. 10, 4147–4150 (May 15, 1971).

Weiss, M. J., Mielczarek, S. R., Kuyatt, C. E., Inelastic electron scattering from nitrous oxide between 5.0 and 20.0 eV energy loss, J. Chem. Phys. **54,** No. 3, 1412–1414 (Feb. 1, 1971).

West, E. D., Data analysis for isoperibol laser calorimetry, Nat. Bur. Stand. (U.S.), Tech. Note 396, 34 pages (Feb. 1971) 40 cents, SD Catalog No. C13.46:396.

White, H. J., Jr., Federal information processing standards index, January 1, 1971, Nat. Bur. Stand. (U.S.), Fed. Info. Process. Stand. Publ. (FIPS Pub) 12, 143 pages, $1.50, SD Catalog No. C13.52:12.

Wiederhorn, S. M., Johnson, H., Effect of pressure on the fracture of glass, J. Appl. Phys. **42,** No. 2, 681–684 (Feb. 1971).

Wiese, W. L., Atomic transition probabilities—A survey of our present knowledge and future needs, (Proc. 2nd. Intern. Conf. Beamfoil Spectroscopy in Nuclear Instruments and Methods, Lysekil, Sweden, June 8–12, 1970). Part II, Lifetimes and transition probabilities, 25–119 in Nucl. Instr. Methods **90,** 25–32 (North-Holland Publ. Co., Amsterdam, The Netherlands, 1970).

Wiese, W. L., Kelleher, D. E., On the cause of the redshifts in white-dwarf spectra, Astrophys. J. 166, L59–L63 (June 1, 1971).

Woelfel, J. B., Paffenbarger, G. C., Expanding and shrinking 7-year-old dentures: report of cases, J. Am. Dental Assoc. **81,** No. 6, 1342–1348 (December 1970).

Wright, J. C., Moos, H. W., Colwell, J. H., Mangum, B. W., Thornton, D. D., $DyPO_4$: A three-dimensional Ising antiferromagnet, Phys. Rev. B. **3,** No. 3, 843–858 (Feb. 1, 1971).

Yakowitz, H., Fiori, C. E., Michaelis, R. E., Standard reference materials: Homogeneity characterization of Fe-3Si alloy, Nat. Bur. Stand. (U.S.), Spec. Publ. 260–22, 30 pages (Feb. 1971) 35 cents, SD Catalog No. C13.10:260–22.

Yates, J. T., Jr., Madey, T. E., Interactions between chemisorbed species: $H_2$ and $N_2$ on (100) tungsten, J. Vacuum Sci. Technol. **8,** No. 1, 63–68 (January–February 1971).

Yokel, F. Y., Dikkers, R. D., Strength of load bearing masonry walls, J. Struct. Div. Am. Soc. Civil Engineers **97,** No. ST5, 1593–1609 (May 1971).

Yokel, F. Y., Mathey, R. G., Dikkers, R. D., Strength of masonry walls under compressive and transverse loads, Nat. Bur. Stand. (U.S.), Bldg. Sci. Ser. **34,** 74 pages (Mar. 1971) 70 cents, SD Catalog No. C13.29/2:24.

Zimmerman, J. E., Recent Developments in superconducting devices, J. Appl. Phys. **42,** No. 1, 30–71 (Jan. 1971).

INDEX SHOULD READ
JANUARY - DECEMBER 1971

JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 75B, Nos. 3 and 4, July–December, 1971

# INDEX TO VOLUME 75B

## July–December 1971


### A

Page

Analytic functions ..... 85
Angular velocity ..... 165
Approximation ..... 85

### B

*Bhat, V. N., Kapoor, S. F.*, The powers of a connected graph are highly hamiltonian ..... 63
Biorthogonal polynomials ..... 1
Block design ..... 23
Boundary value problems ..... 121
Bounded index ..... 141
Bounds to truncation errors in biorthogonal polynomial approximations, with illustrative applications to gamma ray transport distributions ..... 1

### C

Casimir check for the algebraic matrices of the configuration $(d+s)^n p$ ..... 31
Casimir operator ..... 31
Chain rule ..... 97
Checking parameters ..... 31
Combinatorial analysis ..... 31
Commutator ..... 17
Conditions for convergence of solutions ..... 173
Configurations ..... 23
Congruences ..... 67
Contimuum mechanics ..... 97
Convex functions ..... 47
Convex homotopy ..... 47
Convex sets ..... 47
Coupled equations ..... 121

### D

$(d+s)^n p$ ..... 31
Determinant ..... 107
Differential equations ..... 121

### E

Effective axis and angle of rotation ..... 165
Eigenvalues of products ..... 115
Eigenvalues of sums ..... 115
Eigenvalues of sums of hermitian matrices. III ..... 115
Entire function ..... 141
Ergodic source which is not a function of any finite Markov process ..... 133
Error functions ..... 149
Exact solutions ..... 67
Exact solutions of linear equations with rational coefficients ..... 67
Exponential type ..... 141

### F

Fischer inequality ..... 77
*Fong, J. T.*, Some elementary formulas in "matrix calculus" and their applications ..... 79
*Fraenkel, A. S., Loewenthal, D.*, Exact solutions of linear equations with rational coefficients ..... 67

### G

Page

Gamma ray transport ..... 1
*Garbanti, Linda Freede, Thompson, Robert C.*, Eigenvalues of sums of hermitian matrices. III ..... 115
*Geist, Jon*, A new type of boundary value coupling for second order Sturm-Liouville systems ..... 121
*Geller, Murray, Ng, Edward W.*, A table of the integrals of the error function. II. Additions and corrections ..... 149
*Gelman, Harry*, A note on the time dependence of the effective axis and angle of a rotation ..... 165
Generalized matrix function ..... 107
Gradient ..... 79
Graph ..... 63
Gaussian quadrature ..... 85

### H

*Haber, S.*, On certain optimal quadrature formulas ..... 85
Hadamard determinant theorem ..... 77
Hamiltonian ..... 63
Homotopy ..... 47
*Horn, W. A.*, Convex homotopy ..... 47

### I

Improvement of the Fischer inequality ..... 77
Indefinite integrals ..... 149
Inequalities for eigenvalues ..... 115
Information theory ..... 133
Integral equation of first kind ..... 173
Integration ..... 85
Intrinsic vector ..... 165
Invariant factors ..... 81
Invariants ..... 107

### J

*Johnsen, E. C., Storer, T. F.*, A special permutation matrix decomposition for combinatorial design incidence matrices ..... 23

### K

*Kapoor, S. F., Bhat, V. N.*, The powers of connected graph are highly hamiltonian ..... 63
Kinematics of a rigid body ..... 165
Konig's theorem ..... 23

### L

Linear equations ..... 67
Linear transformations on matrices ..... 107
*Loewenthal, D., Fraenkel, A. S.*, Exact solutions of linear equations with rational coefficients ..... 67

### M

*Marcus, Marvin*, Linear transformations on matrices ..... 107
Markov process ..... 133
Matrices ..... 23, 79, 107
0,1 Matrices ..... 23
Matrix ..... 17
Matrix calculus ..... 79

Page

Matrix equations ..... 23
Maximum modulus ..... 141
*Mazur, Jacob*, A numerical solution of integral equation of the first kind applied to slit correction in small-angle x-ray scattering ..... 173
*Merris, R.*, An improvement of the Fischer inequality ..... 77
Modular arithmetic ..... 67
Moment method ..... 1

N

Neutron transport ..... 1
*Newman, M.*, On the Smith normal form ..... 81
New type of boundary value coupling for second order Sturm-Liouville systems ..... 121
*Ng, Edward W., Geller, Murray*, A table of the integrals of the error function. Additions and corrections ..... 149
Note on the time dependence of the effective axis and angle of a rotation ..... 165
Numerical solution of integral equation of the first kind applied to slit correction in small-angle x-ray scattering ..... 173

O

On certain optimal quadrature formulas ..... 85
On entire functions of exponential type ..... 141
Orthogonal ..... 17

P

Partial differentiation ..... 79
Permanent ..... 77
Permutation matrix decomposition ..... 23
Polynomial truncation errors ..... 1
Positive definite hermitian matrix ..... 77
Power of a graph ..... 63
Powers of a connected graph are highly hamiltonian ..... 63
Principal ideal rings ..... 81
Probability theory ..... 133
Product rule ..... 79

Q

Quadrature ..... 85

R

Rank ..... 107
Ring theory ..... 133

Page

Rotation group ..... 165
*Roth, C.*, The Casimir check for the algebraic matrices of the configuration $(d+s)^n p$ ..... 31

S

Self-adjoint equations ..... 121
*Shah, S. M., Sisarcick, W. C.*, On entire functions of exponential type ..... 141
Shifting operators ..... 173
Sievert function ..... 1
Singular values ..... 115
*Sisarcick, W. C., Shah, S. M.*, On entire functions of exponential type ..... 141
Skew-symmetric ..... 17
Slit correction ..... 173
Slit wieghting function ..... 173
Small-angle x-ray scattering ..... 173
*Smith, J. M.*, A theorem on matrix commutators ..... 17
Smith normal form ..... 81
Some elementary formulas in "matrix calculus" and their applications ..... 79
Spatial rotation ..... 165
Special functions ..... 149
Special permutation matrix decomposition for combinatorial design incidence matrices ..... 23
*Spencer, L. V.*, Bounds to truncation errors in biorthogonal polynomial approximations, with illustrative applications to gamma ray transport distribution ..... 1
Stochastic process ..... 133
*Storer, T. F., Johnsen, E. C.*, A special permutation matrix decomposition for combinatorial design incidence matrices ..... 23
Strong numerical instability ..... 173
Sturm-Liouville systems ..... 121
$SU_3$ ..... 31

T

Table of integrals of the error function. II. Additions and corrections ..... 149
Tensor function ..... 79
Theorem on matrix commutators ..... 17
*Thompson, Robert C., Garbanti, Linda Freede*, Eigenvalues of sums of hermitian matrices. III ..... 115
Trace ..... 79
*Traub, R. D.*, An ergodic source which is not a function of any finite Markov process ..... 133




U. S. GOVERNMENT PRINTING OFFICE : 1972 OL - 449-764

UNITED STATES
GOVERNMENT PRINTING OFFICE
PUBLIC DOCUMENTS DEPARTMENT
WASHINGTON, D.C. 20402

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. GOVERNMENT PRINTING OFFICE

U.S.MAIL

JOURNAL OF RESEARCH of the National Bureau of Standards Vol. 75B, Nos. 3&4 July–December 1971
B. Mathematical Sciences

## Contents

| | Page |
|---|---|
| **Some elementary formulas in "matrix calculus" and their applications. Jeffrey T. Fong.** | **97** |
| **Linear transformations on matrices. Marvin Marcus.** | **107** |
| **Eigenvalues of sums of hermitian matrices. III. Robert C. Thompson and Linda Freede Garbanti.** | **115** |
| **A new type of boundary value coupling for second order Sturm-Liouville systems. Jon Geist.** | **121** |
| **An ergodic source which is not a function of any finite Markov process. Roger D. Traub.** | **133** |
| **On entire functions of exponential type. S. M. Shah and W. C. Sisarcick.** | **141** |
| **A table of integrals of the error function. II. Additions and corrections. Murray Geller and Edward W. Ng.** | **149** |
| **A note on the time dependence of the effective axis and angle of a rotation. Harry Gelman.** | **165** |
| **A numerical solution of integral equation of the first kind applied to slit correction in small-angle x-ray scattering. Jacob Mazur.** | **173** |
| **Publications of the National Bureau of Standards.** | **189** |
| **Index to volume 75B, January to December 1971.** | **195** |

Library of Congress Catalog Card Number: 62–4414

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. (Order by SD Catalog No. C 13.22/sec.B:75B (3&4). Price $1.25 (single copy). Subscription price: $5.00 a year; $1.25 additional for foreign mailing.

UNITED STATES GOVERNMENT PRINTING OFFICE, WASHINGTON : 1971

7
3